Acid-base thermochemistry of gaseous oxygen and sulfur substituted amino acids (Ser, Thr, Cys, Met)

Vanessa Riffet, Gilles Frison, Guy Bouchoux

PCCP

View Online
Dynamic Article Links

Cite this: *Phys. Chem. Chem. Phys.*, 2011, **13**, 18561–18580

www.rsc.org/pccp

PAPER

Acid–base thermochemistry of gaseous oxygen and sulfur substituted amino acids (Ser, Thr, Cys, Met)†

Vanessa Riffet, Gilles Frison and Guy Bouchoux*

Received 5th July 2011, Accepted 26th August 2011
DOI: 10.1039/c1cp22206f

Acid–base thermochemistry of isolated amino acids containing oxygen or sulfur in their side chain (serine, threonine, cysteine and methionine) have been examined by quantum chemical computations. Density functional theory (DFT) was used, with B3LYP, B97-D and M06-2X functionals using the 6-31+G(d,p) basis set for geometry optimizations and the larger 6-311++G(3df,2p) basis set for energy computations. Composite methods CBS-QB3, G3B3, G4MP2 and G4 were applied to large sets of neutral, protonated and deprotonated conformers. Conformational analysis of these species, based on chemical approach and AMOEBA force field calculations, has been used to identify the lowest energy conformers and to estimate the population of conformers expected to be present at thermal equilibrium at 298 K. It is observed that G4, G4MP2, G3B3, CBS-QB3 composite methods and M06-2X DFT lead to similar conformer energies. Thermochemical parameters have been computed using either the most stable conformers or equilibrium populations of conformers. Comparison of experimental and theoretical proton affinities and $\Delta_{acid}H$ shows that the G4 method provides the better agreement with deviations of less than 1.5 kJ mol^{-1}. From this point of view, a set of evaluated thermochemical quantities for serine, threonine, cysteine and methionine may be proposed: PA = 912, 919, 903, 938; GB = 878, 886, 870, 899; $\Delta_{acid}H$ = 1393, 1391, 1396, 1411; $\Delta_{acid}G$ = 1363, 1362, 1367, 1382 kJ mol^{-1}. This study also confirms that a non-negligible ΔpS° is associated with protonation of methionine and that the most acidic hydrogen of cysteine in the gas phase is that of the SH group. In several instances new conformers were identified thus suggesting a re-examination of several IRMPD spectra.

1. Introduction

Detailed knowledge on the structural and energetic properties of isolated neutral, deprotonated and protonated amino acids is of interest to numerous areas of chemistry and biochemistry. It is, for example, well established that proton transfer involving amino acids and their polymers plays an important role in biological processes. From another point of view, structure elucidation of amino acids and their polymers by mass spectrometry results, most of the time, from the observation and dissociation of deprotonated or protonated species. The intrinsic acidity and basicity of isolated amino acids provide essential clues for the understanding of these processes.

Thermochemical data associated with protonation (proton affinity, PA, and gas phase basicity, GB) or deprotonation ($\Delta_{acid}H$ and $\Delta_{acid}G$) of molecular species in the gas phase are defined by reactions (1) and (2).

$$\underset{(gas)}{MH^+} \rightarrow \underset{(gas)}{M} + \underset{(gas)}{H^+} \qquad (1)$$

$$PA(M) = \Delta_1 H^{o} \text{ and } GB(M) = \Delta_1 G^{o}$$

$$\underset{(gas)}{M} \rightarrow \underset{(gas)}{[M-H]^-} + \underset{(gas)}{H^+} \qquad (2)$$

$$\Delta_{acid}H(M) = \Delta_2 H^{o} \text{ and } \Delta_{acid}G(M) = \Delta_2 G^{o}$$

These quantities can be obtained experimentally from mass spectrometry measurements,[1–3] and theoretically, from quantum

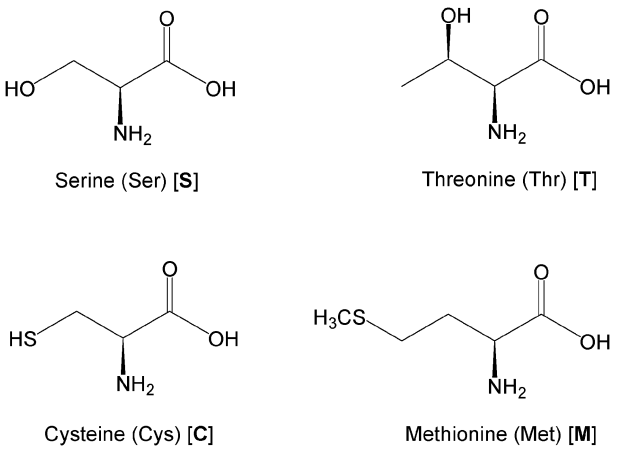
Laboratoire des Mécanismes Réactionnels, UMR CNRS 7651, Département de Chimie, Ecole Polytechnique, F-91128 Palaiseau cedex, France. E-mail: bouchoux@dcmr.polytechnique.fr; Fax: 33 1 69 33 48 03; Tel: 33 1 69 33 48 42

† Electronic supplementary information (ESI) available: enthalpies, entropies and free energies of the most stable conformers of neutral, protonated and deprotonated serine, threonine, cysteine and methionine computed at 298 K using the B3LYP, B97-D and M06-2X functionals. The CBS-QB3, G3B3, G4MP2 and G4 composite methods are given in Tables T1 to T4. Relative H^{o}_{298} and G^{o}_{298} energies and optimized B3LYP/6-31+G(d,p) geometries of all neutral, protonated and deprotonated species are presented in Fig. S1 to S21. Signed deviations (method-G4 and method-experiment) on the computed deprotonation and protonation thermochemistry of the studied α-amino acids are presented in Tables T5 to T8. Computed thermochemical data of reference compounds ammonia and benzoic acid are given in Table T9. The correspondence between conformers' names in excel tables and charts are contained in Table T10. See DOI: 10.1039/c1cp22206f

chemistry computations.[4,5] Moreover, the latter methods bring valuable molecular structure information on the neutral and ionized species hardly obtainable experimentally. A crucial point in the computation of thermochemical quantities concerning amino acids, or, more generally, molecular polyfunctional systems, is the correct consideration of non-covalent interactions. Accordingly, in these systems, the stability of a given structure results from equilibrium between different non-covalent interactions such as hydrogen bonds, salt bridges and van der Waals interactions that include London dispersion forces. Non-covalent interactions consequently influence the conformational population of these molecules and their protonated or deprotonated forms. Description of long-range interactions in terms of geometry and energy has been a challenge of theoretical chemistry during the latest years.[6] It has long been recognized that the correlation energy of the system plays an essential part in its stabilization energy. Møller–Plesset theory and coupled cluster method are the usual post Hartree–Fock treatments allowing consideration of such an effect. Good illustrations are composite methods which have been developed with the aim to offer "chemical" accuracy (*i.e.* $\sim$5 kJ mol^{-1}) on enthalpic quantities.[7] Other attractive approaches involve the use of density functional theory (DFT) since it is less computationally demanding. However, the most widely used B3LYP method seems to present several shortcomings.[8] Recently, several DFT methods including explicit treatment of dispersion effects were introduced in order to provide a more confident description of noncovalent interactions.[8,9] It was consequently of interest to test and to compare these methods on typical systems where sufficient experimental data are available.

The goal of the present study is to evaluate accurately the acid–base thermochemistry of the series of naturally occurring α-amino acids bearing an oxygen or a sulfur atom on their side chain, namely, L-serine(2*S*), L-threonine(2*S*,3*R*), L-cysteine(2*R*) and L-methionine(2*S*) (Scheme 1).

In the first step, starting geometries were generated through systematic manual conformational exploration and molecular dynamics and Monte-Carlo calculations with the AMOEBA force field.[10] The resulting geometries of neutral (M) and ionized (MH^+ and $[M-H]^-$) amino acids have been subsequently optimized with quantum chemical tools in order to identify the most stable conformers. For this purpose, we



Scheme 1

select three types of density functional: firstly, the popular hybrid B3LYP functional[11] since it is presently widely used as a computational tool by many researchers, secondly, the pure semi-local generalized gradient correlation (GGA) B97-D[9] functionals where long-range dispersion effects are included through a R^{-6} classical potential, and, thirdly, the meta-GGA M06-2X functional[8,12] which includes a high percentage of Hartree–Fock exchanges and a parameterization based on various sets of reference data including *inter alia* interaction energies of non-covalent complexes. The reasons that prompt us to use this latter functional are that it is claimed to best describe hydrogen bonded interactions and to provide accurate thermodynamic data such as barrier heights and isomerization energies.[8,12,13] Moreover, M06-2X has also been demonstrated to reproduce nicely conformer energies and populations in systems dominated by non-covalent and electrostatic interactions.[14]

Conformer relative energies, protonation and deprotonation thermochemistry were then computed using a panel of composite methods which offer, at the present time, the best accuracy on energetic quantities. The four selected methods, namely CBS-QB3, G3B3, G4MP2 and G4,[15–18] have been assessed on a test set containing up to 454 energies and show average deviation from experiment generally below $\sim$5 kJ mol^{-1}. A complete re-examination of the available experimental gas phase basicity and acidity data of serine, threonine, cysteine and methionine is also presented in order to allow a meaningful comparison with theory and to propose newly evaluated thermochemical data.

2. Methodology

2.1. Quantum chemistry

Molecular orbital calculations have been conducted using the GAUSSIAN09 suite of programs.[19] In order to achieve a thorough search of the most stable conformers, two strategies have been used since recent work showed that a single strategy does not always allow location of all conformers when non-covalent interactions are operating.[20] Our first strategy was a "chemical approach" which consisted in exploring the potential energy surface by systematically constructing the various staggered conformers by rotation around the CC, CN or CO bonds. For each amino acid, this corresponds to one to several hundreds of trial conformations for each neutral, protonated and deprotonated forms. In particular in the case of methionine, where the number of internal degrees of freedom becomes important, the conformational landscapes were also explored using both molecular dynamics and Monte Carlo simulations as well as a basin-hopping conformational scanning algorithm as implemented in the Tinker molecular modeling package.[21] For this second strategy, we use the polarizable AMOEBA force field[10] which gives generally very good agreement for the determination of geometrical structure and relative conformational energies.[20,22] The parameter set has been used as included in the Tinker package when available. As done previously for protonated glycine,[20] the parameters for the neutral COOH and NH_2 terminations were adjusted based on those of the carboxylic group of the aspartic acid and

the methylamine, respectively, which are already available in the AMOEBA parameter set. After having defined the starting geometries of the various conformers by these prior procedures, geometries were optimized at a DFT level (*i.e.* B3LYP, B97-D and M06-2X) using the double-ζ quality basis set 6-31+G(d,p). This procedure furnishes then 74, 13 and 11 unique conformers for neutral, protonated and deprotonated serine, respectively (65, 12 and 11 for Thr; 87, 21 and 39 for Cys, and 32, 6, 24 for Met, but, in this latter case, in a limited 15 kJ mol^{-1} G°_{298} range). The most stable conformers located at these levels in a 15 kJ mol^{-1} range were then subjected to more elaborate computations using the composite methods CBS-QB3, G3B3, G4MP2 and G4. Vibrational frequencies were calculated at the same level as the geometry optimization to confirm that the structures are true minima on the potential energy surfaces. Single point energy calculations were performed at the 6-31+G(d,p) optimized geometries using triple-ζ quality basis set 6-311++G(3df,2p) on several key conformers.

Composite CBS-QB3,[15] G3B3,[16] G4MP2[17] and G4[18] methods are based on a series of quantum chemistry calculations combined assuming additivity of the energy terms. The complete basis set, CBS-QB3 model[15] uses optimized geometry and frequencies calculated at the B3LYP/6-311G(2d,p) level. Then, MP2/6-311+G(2df,2p) energy and CBS extrapolation are computed. Further, MP4(SDQ)/6-31+G(d,p) and QCISD(T)/6-311G(d,p) single points are performed and finally, two empiric correction terms including effects of absolute overlap integral and spin contamination, are considered in the overall energy estimate.

The G3B3 methods[16] use geometry optimized at the B3LYP/6-31G(d) level, zero point vibrational energy is obtained from vibrational frequencies calculated at this level and scaled by a factor 0.96. In a second step, single point energy calculations are performed using (i) the frozen core QCISD(T)/6-31G(d) approximation, (ii) MP2(full) computation using the G3Large basis set (roughly a 6-311+G(2df,2p) basis for the first row elements) set and (iii) MP4(FC) computations using the 6-31G(d) and larger basis sets such as 6-31+G(d) and 6-31G(2df,p). Finally a "high level correction" (HLC) is introduced to account for remaining deficiencies in the energy computation. The G4 technique[18] differs from G3B3 by (i) equilibrium structure and harmonic frequencies (scaled by 0.9854) calculated at the B3LYP/6-31G(2df,p) level, (ii) calculation of the Hartree–Fock energy limit, (iii) single point correlation energies are calculated at the CCSD(T)/6-31G(d) rather than the QCISD(T) /6-31G(d) level, (iv) MP2(full) computation uses the G3LargeXP basis set which consists of the G3Large basis improved by adding new d-polarization functions, and finally two additional parameters (A′ and E) are added in the HLC correction. In the G4(MP2) method,[17] MP3 and MP4 large basis set calculations are eliminated from the G4 procedure.

2.2. Thermochemistry

As recalled in the Introduction, the thermochemical quantities associated with the intrinsic basicity and acidity of a molecule M are defined as the standard enthalpy and standard Gibbs free energy of reactions (1) and (2), respectively.[3] Starting from these definitions, PA(M) and $\Delta_{acid}H$(M) at temperature T, may be calculated using:

$$PA_T(M) = \Delta_1 H^{\circ}_T = H^{\circ}_T(M) + H^{\circ}_T(H^+) - H^{\circ}_T(MH^+) \quad (3)$$

$$\Delta_{acid}H_T(M) = \Delta_2 H^{\circ}_T = H^{\circ}_T([M - H]^-) + H^{\circ}_T(H^+) - H^{\circ}_T(M) \quad (4)$$

where H°_T represents the enthalpy of each species.

Since PA_T(M) and $\Delta_{acid}H$(M) are molar quantities, their estimate should in principle involve one mol of M and MH^+ (or $[M–H]^-$) in thermal equilibrium at temperature T. The usual computational approach assumes that each mol of reactant and product of reactions (1) or (2) contains only the most stable conformers of each species. Under these circumstances, eqn (3) and (4) may be directly used to derive "most stable conformers" (msc) quantities, $PA_T(M)_{msc}$ and $\Delta_{acid}H(M)_{msc}$. At 298 K these equations reduce to eqn (5) and (6):

$$PA(M)_{msc} = H^{\circ}_{298}(M)_{msc} - H^{\circ}_{298}(MH^+)_{msc} + 6.19 \text{ kJ mol}^{-1} \quad (5)$$

$$\Delta_{acid}H(M)_{msc} = H^{\circ}_{298}([M - H]^-)_{msc} - H^{\circ}_{298}(M)_{msc} + 6.19 \text{ kJ mol}^{-1} \quad (6)$$

since the $H^{\circ}_T(H^+)$ term is simply equal to 5/2 RT.

Rigorously, however, a distribution of conformers should be considered for both reactants and products. Assuming thermal equilibrium, a Boltzmann distribution may be used to derive the molar fractions x_i of each conformer at temperature T using eqn (7):

$$x_i = \exp(-G_i/RT)/\sum_1^N \exp(-G_i/RT) \quad (7)$$

where N is the total number of conformers and G_i their individual Gibbs free energies.

Averaged enthalpy value of one mole of mixture may be determined from the enthalpies of the separate species according to:

$$[H^{\circ}_T]_{average} = \sum_1^N x_i (H^{\circ}_T)_i \quad (8)$$

thus leading to:

$$PA(M)_{average} = [H^{\circ}_{298}(M)]_{average} - [H^{\circ}_{298}(MH^+)]_{average} + 6.19 \text{ kJ mol}^{-1} \quad (9)$$

$$\Delta_{acid}H(M)_{average} = [H^{\circ}_{298}([M - H]^-)]_{average} - [H^{\circ}_{298}(M)]_{average} + 6.19 \text{ kJ mol}^{-1} \quad (10)$$

Similarly, the "most stable conformer" and "averaged" definitions may be introduced for the gas phase basicities, GB(M) and for the corresponding acidity parameter $\Delta_{acid}G$(M). It should be noted however that the averaged entropy of a collection of conformers should include the entropy of mixing and be consequently estimated *via* eqn (11):

$$[S^{\circ}_T]_{average} = \sum_1^N x_i (S^{\circ}_T)_i - R\sum_1^N x_i \ln x_i \quad (11)$$

where the second component corresponds to the mixing contribution.

2.3. Nomenclature

Nomenclature used to describe the various conformers parallels that used previously in a preceding study of aliphatic amino acids.[23] Briefly, three types of internal hydrogen bonds located in the C(H,R)NH_2COOH framework have been defined as **I**, **II** and **III** (Chart 1). The three possible orientations of the amino group inside the amino acid moiety are designated by letters **A**, **B** or **C** (Chart 1). Note that for type **I** conformers, a secondary hydrogen bond O(S)H$\cdots$$NH_2$ is generally occurring. The denomination **I′** will be used to describe the situation where $NH_2\cdots$O(S)H. Conformers of types **I** and **III** discussed in the text of the present study are invariably associated with C–N rotamers **C**. For the purpose of clarity, the **C** letter is consequently not indicated. Concerning the conformers generated by rotation around the C(2)–C(3) bond, the nomenclature *gauche* (abbreviated *g*+ or *g*−) and *antiperiplanar* (abbreviated *a*) has been used as noted in Chart 2.

Most stable protonated amino acids conformers are characterized by two general types of arrangements where the internal hydrogen bonding is established either toward the carbonyl or the hydroxyl oxygen. These two situations are denoted **HI** and **HII** as defined in Chart 3. In the case of deprotonated amino acids the two possible orientations of the NH_2 group with respect to the carboxylate moiety are denoted **-H_A** and **-H_B** (Chart 3) by analogy with the nomenclature used to describe the neutral molecules (Chart 1).

3. Results and discussion

In the following section we will present first the essential results of our extensive exploration of the conformational space of neutral and ionized amino acids Ser, Thr, Cys and Met. Detailed data are given in the electronic supplementary information.† In the second and third parts, protonation and deprotonation thermochemistry are examined and compared with experimental data.

3.1 Conformational landscape of neutral, protonated and deprotonated amino acids

Serine. Previous explorations of the conformational space of neutral serine were performed using Hartree–Fock,[24,25] B3LYP,[26–31] MP2[27b] and G2(MP2)[32] or G3(MP2)[33] methods

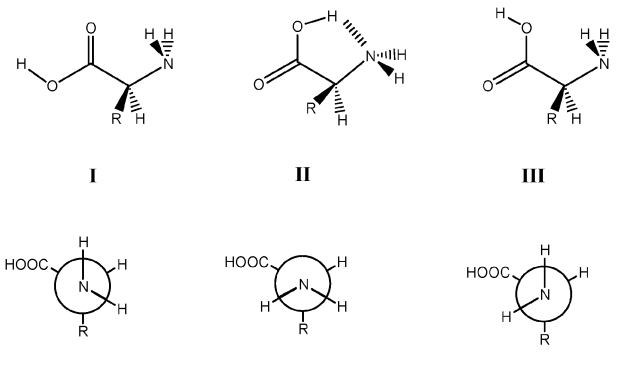
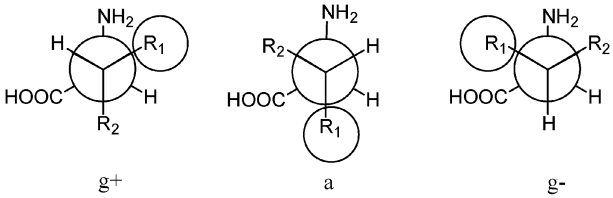


Chart 1



R_1=OH, SH, CH_2SCH_3; R_2=H, CH_3

Chart 2

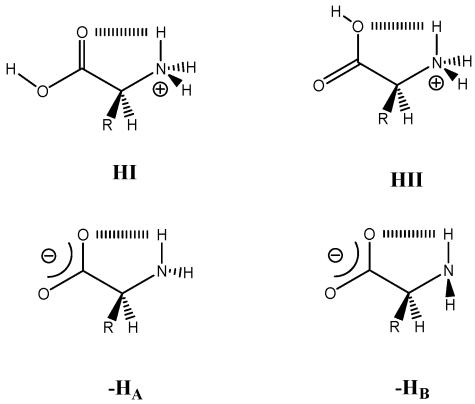


Chart 3

(note that enantiomer 2*R* rather than 2*S* was considered in ref. 24 and 26). Depending upon the theoretical level and the strategy used to explore the potential energy surface, different conformers were considered as the most stable. In the present study, our conformational search followed by geometry optimization at the B97-D, B3LYP and M06-2X levels using the 6-31+G(d,p) basis set leads to 74 stable forms among which the most stable have been further examined by the composite methods CBS-QB3, G3B3, G4MP2 and G4. Only the results concerning the eighth most stable conformers will be discussed here (Fig. 1 and 2), the full set of data is available in Table T1 and Fig. S1 and S2 of electronic supplementary information (ESI).† Fig. 1 illustrates the influence of the theoretical level of theory on the relative stability (H°_{298} and G°_{298}) of the first eight conformers. It clearly appears that the four composite methods predict similar relative H°_{298} and G°_{298} energies. The situation is different for DFT methods. If M06-2X behaves similarly to the composite methods, the relative energy orders given by B97-D and B3LYP seriously disagree. A second observation is that the relative stability orders given by H°_{298} and G°_{298} are different, obviously because of different third law entropies between conformers. In order to understand this effect it is necessary to examine the geometrical characteristics of these conformers (Fig. 2).

Conformers **SI_Cg+**, **SI_Ag−**, **SI'_Cg−** present a $NH_2\cdots$O=C–OH(*syn*), (type **I**), arrangement while **$SIII_Cg$+** and **$SIII_Cg$−** correspond to the same kind of internal hydrogen bond but involve the less basic hydroxyl oxygen as acceptor: $NH_2\cdots$OH–C=O(*syn*), (type **III**). Note that conformers **SI_Cg+** and **SI_Ag−** present a supplementary source of stabilization with a $CH_2OH\cdots NH_2$ hydrogen bond while conformer **SI'_Cg−** is characterized by a reverse $NH_2\cdots O(H)CH_2$ interaction.

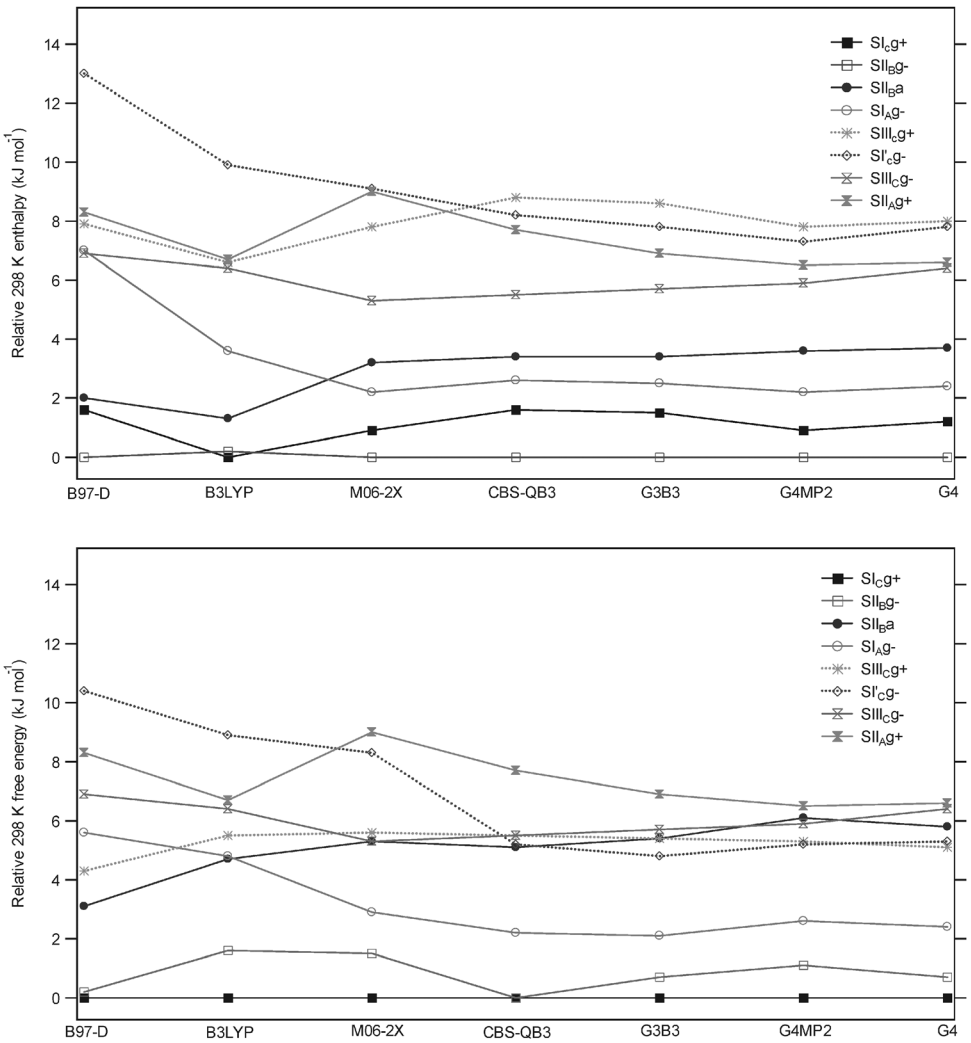


Fig. 1 Relative H°_{298} and G°_{298} (kJ mol^{-1}) of the eight most stable conformers of neutral L-serine.

Finally, the three conformers **SII$_{Bg}$**−, **SII$_{Ba}$** and **SII$_{Ag}$**+ correspond to an *anti* conformation of the acidic moiety thus allowing $H_2N\cdots HO$–C=O(*anti*), (type **II**) internal hydrogen bonding. It is known that the H-bond interaction involved in type **II** conformers is particularly efficient since it involves a strong acidic hydrogen and a strong basic nitrogen atom.[23,34] This is attested to by the short distances between the involved atoms which lie around 1.9 Å (see **SII$_{Bg}$**−, **SII$_{Ba}$** and **SII$_{Ag}$**+, Fig. 2). By contrast, the H-bond distance is situated between 2.2 and 2.5 Å for conformers of types **I** or **III**. Comparable observation can be made with the hydrogen bonds involving the hydrogen of the alcohol moiety and either the amino nitrogen (conformers **SI$_{Cg}$**+, **SI$_{Ag}$**−, **SIII$_{Cg}$**+) or the oxygen of the carbonyl group (conformers **SII$_{Ba}$**, **SIII$_{Cg}$**−). From this point of view, the case of conformer **SII$_{Ba}$** should be underlined since it presents the shortest $CH_2OH\cdots NH_2$ distance (1.978 Å, Fig. 2). These differences in internal H-bonding induce differences in rotational barriers and, consequently, in entropies. It is thus not surprising to observe that conformer **SII$_{Ba}$** has the lowest S°_{298} (349 J mol^{-1} K^{-1}), followed by conformer **SII$_{Bg}$**− (S°_{298} = 353 J mol^{-1} K^{-1}), whereas conformers **SIII$_{Cg}$**+ and **SI′$_{Cg}$**− present S°_{298} close to 365 J mol^{-1} K^{-1}. This results in the observed change in stability order between H°_{298} (**SII$_{Bg}$**− < **SI$_{Cg}$**+ < **SI$_{Ag}$**− < **SII$_{Ba}$** < **SIII$_{Cg}$**−, **SII$_{Ag}$**+ < **SIII$_{Cg}$**+, **SI′$_{Cg}$**−) and G°_{298} (**SI$_{Cg}$**+ < **SII$_{Bg}$**− < **SI$_{Ag}$**− < **SIII$_{Cg}$**+, **SI′$_{Cg}$**− < **SII$_{Ba}$** < **SII$_{Ag}$**+, **SIII$_{Cg}$**−). Using the G4 relative free energies of the eight more stable conformers of neutral serine, we can calculate the conformer population at 298 K, **SI$_{Cg}$**+/**SII$_{Bg}$**−/**SI$_{Ag}$**−/**SIII$_{Cg}$**+/**SI′$_{Cg}$**−/**SII$_{Ba}$**/**SII$_{Ag}$**+/**SIII$_{Cg}$**−: 38.1/28.4/14.4/4.9/4.6/3.8/3.1/2.7%.

It is satisfactory to observe that the infrared spectrum of neutral serine, isolated in a low temperature argon matrix,[27] has been interpreted by the occurrence of a mixture of conformers **SI$_{Cg}$**+, **SII$_{Bg}$**−, **SII$_{Ba}$** and **SI$_{Ag}$**−. A more recent microwave study of gaseous serine produced by laser ablation has been interpreted by the existence of a population of conformers **SI$_{Cg}$**+, **SII$_{Bg}$**−, **SI$_{Ag}$**−, **SII$_{Ba}$**, **SIII$_{Cg}$**− and **SII$_{Ag}$**+.[27b] These conclusions are in correct agreement with our computational results.

13 unique structures have been identified for protonated serine after geometry optimizations at the B97-D, B3LYP and M06-2X levels using the 6-31+G(d,p) basis set. Enthalpies and free energies at 298 K were then computed using the selected

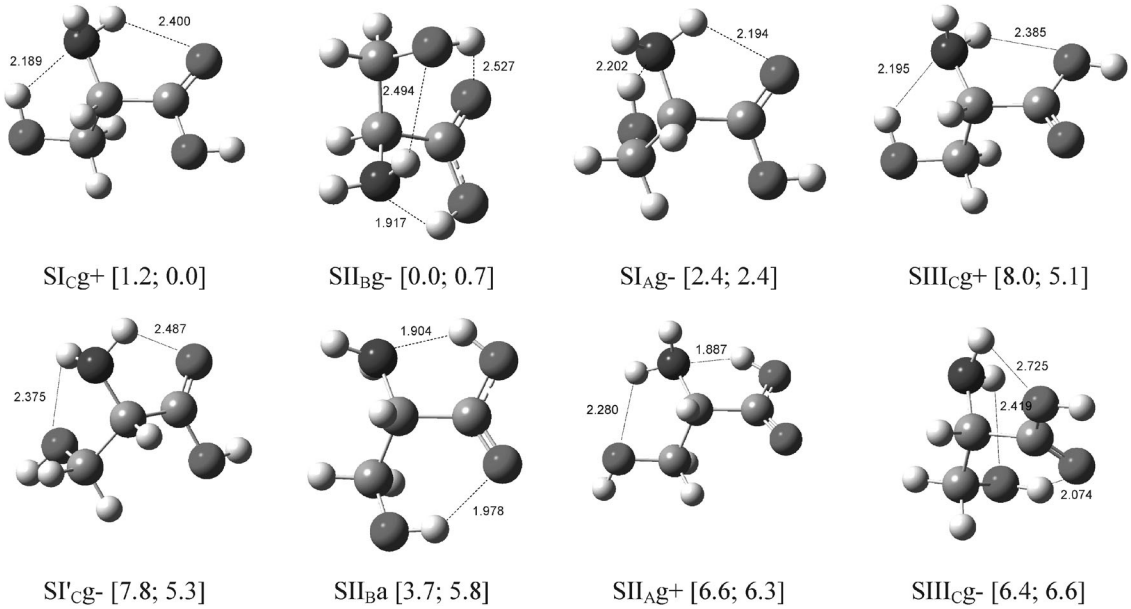


Fig. 2 Optimized geometries (B3LYP/6-31G(2df,p)) of the eight most stable conformers of neutral L-serine (in brackets, relative H°_{298} and G°_{298} in kJ mol^{-1} calculated at the G4 level).

composite methods (Table T1, Fig. S3 and S4 of ESI†). Results are illustrated by Fig. 3 for the four most stable conformers **SHIg**−, **SHIg**+, **SHIIg**− and **SHIIg**+. As already observed for neutral serine, DFT methods behave differently from the composite methods. A clear increase in the H°_{298} differences between **SHIg**− and the three other conformers is observed when passing from B97-D to CBS-QB3. By contrast, these differences remain constant between CBS-QB3 and the more sophisticated G4 method. Evolution of the free energies at 298 K follows exactly the same trends, with however a slight compression of the energy scale. As for neutral serine, this difference may again be explained by considering the entropies of these four conformers, consideration which may be enlightened by looking at their geometrical features (Fig. 4).

The four structures **SHIg**−, **SHIg**+, **SHIIg**− and **SHIIg**+ present a *syn* HOCO arrangement of the acidic moiety allowing the establishment of a strong hydrogen bond with one of the hydrogens of the NH_3^+ group. The most favorable interaction obviously occurs with the oxygen of the carbonyl group thus leading to the two most stable conformers **SHIg**− and **SHIg**+ (upper part of Fig. 4). The other possibility *i.e.* interaction between NH_3^+ and the oxygen of the hydroxyl group is less favorable and leads to conformers **SHIIg**− and **SHIIg**+ situated ~15 kJ mol^{-1} above **SHIg**− and **SHIg**+. The most significant structural difference between **SHIg**− and **SHIg**+ (or **SHIIg**− and **SHIIg**+) lies in the C=O···H$_3$N$^+$ distance. This distance is shorter for conformers **g**− thus explaining why they are more stable than their counterparts **g**+. The second consequence is that the third law entropies of **SHIg**− and **SHIIg**− are lower than those of **SHIg**+ and **SHIIg**+, respectively. As observed in Fig. 3, this entropy effect produces a decrease in the G°_{298} difference, with respect to H°_{298}, between the couples **SHIg**− and **SHIg**+ on one hand and **SHIIg**− and **SHIIg**+ on the other.

It is evident from the present data that the two conformers **SHIg**− and **SHIg**+ will describe most of the population of protonated serine at 298 K. Indeed, using the G°_{298} calculated at the G4 level, a **SHIg**−/**SHIg**+/**SHIIg**−/**SHIIg**+ ratio equal to 68.3/31.1/0.4/0.2% is estimated. The largest stability of **SHIg**− observed both in enthalpy and in free energy in the present study up to the G4 level confirms earlier conclusions based on less accurate calculations.[26,30,32,33] Experimentally, the IRMPD spectrum of protonated serine presents three absorption bands at 1158 cm^{-1}, 1460 cm^{-1} and 1794 cm^{-1}.[35] This was attributed to a HOCO bend, a NH_3 umbrella and CO stretching, respectively for conformer **SHIg**− (SH01 in the original paper).[35] Computation at the B3LYP/6-31G(2df,p) level leads to (unscaled) theoretical vibrational frequencies of 1194, 1467 ± 4 and 1837 cm^{-1} for both conformers **SHIg**− and **SHIg**+ thus leading to the conclusion that the presence of the two conformers may also account for the experimental observation.

Geometry optimization of the 162 trial structures for deprotonated serine [Ser−H]$^-$ at the DFT levels converged on 11 unique conformers (Table T1 and Fig. S5 and S6 in ESI†). The relative enthalpies and free energies at 298 K of the four most stable conformers **S-H$_B$a**, **S-H$_A$a**, **S-H$_B$g**− and **S-H$_A$g**+ are reproduced in Fig. 5 as a function of the theoretical method, their optimized geometries are presented in the lower part of Fig. 4. As shown in Fig. 5, conformers **S-H$_A$a** and **S-H$_B$a**, are almost of identical stabilities at the composite CBS-QB3, G3B3, G4MP2 and G4 levels. As previously observed for neutral and protonated species, the DFT methods present variable results. Discrepancies with the composite methods are however limited to a few kJ mol^{-1}. Differences in H°_{298} and G°_{298} relative values are less pronounced here, compared with the data obtained for neutral and protonated serine. Accordingly, entropies are close together for the four conformers: 345 J mol^{-1} K^{-1} for **S-H$_A$a** and **S-H$_B$a**, 348 and 351 J mol^{-1} K^{-1} for **S-H$_B$g**− and **S-H$_A$g**+ in accordance with structural similarities. In fact, the three most stable structures **S-H$_B$a**, **S-H$_A$a** and **S-H$_B$g**− are characterized by strong intramolecular CH$_2$OH···OCO− interactions

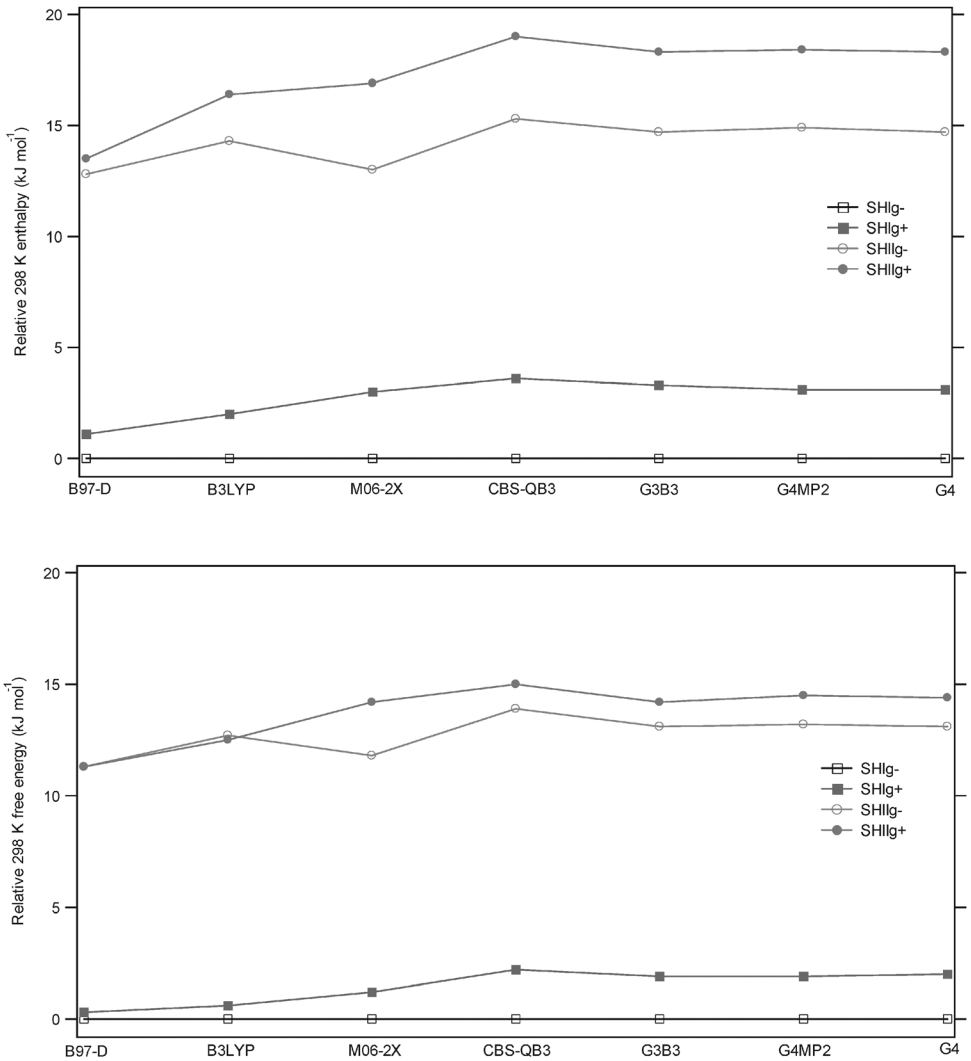


Fig. 3 Relative H°_{298} and G°_{298} (kJ mol^{-1}) of the four most stable conformers of protonated L-serine.

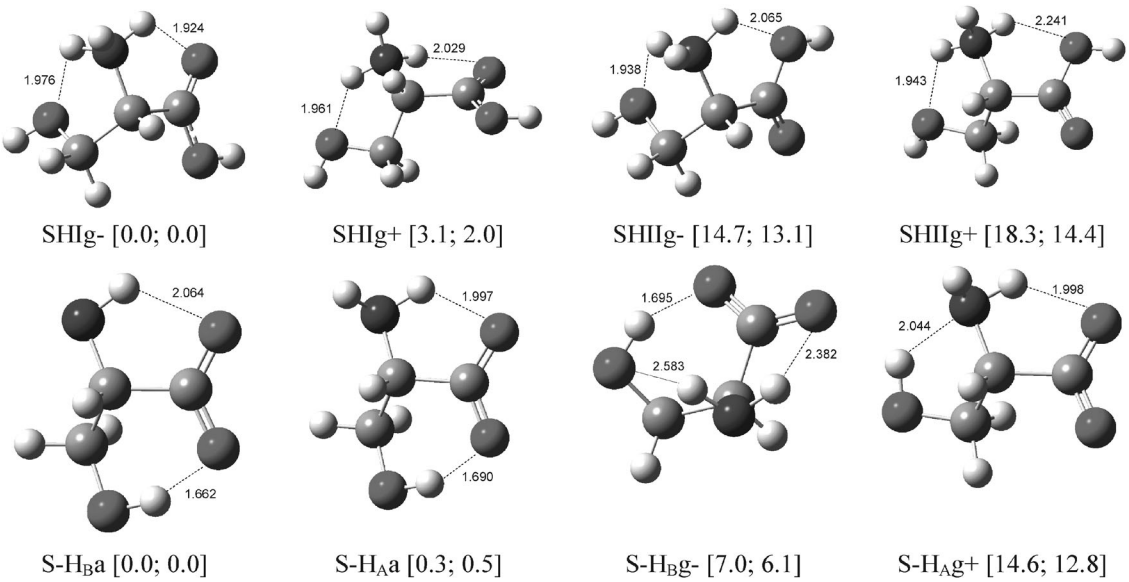


Fig. 4 Optimized geometries (B3LYP/6-31G(2df,p)) of the four most stable conformers of protonated and deprotonated L-serine (in brackets, relative H°_{298} and G°_{298} in kJ mol^{-1} calculated at the G4 level).

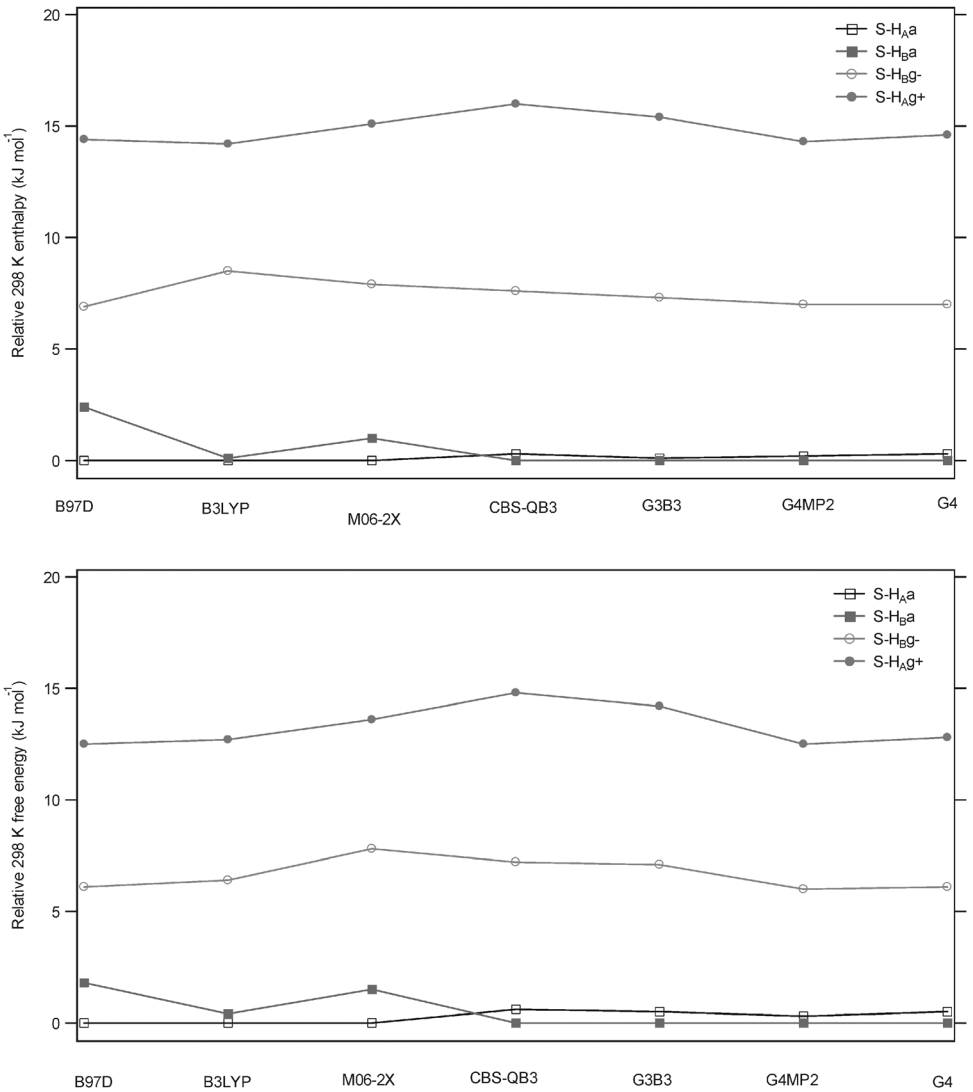


Fig. 5 Relative H°_{298} and G°_{298} (kJ mol^{-1}) of the four most stable conformers of deprotonated L-serine.

(corresponding to the shorter distances of 1.662 to 1.695 Å, Fig. 4). The essential structural differences between **S-H$_A$a** and **S-H$_B$a** arise from the orientation of the NH_2 group, the hydrogen of which being in favorable interaction with the second oxygen of the COO^- moiety. Conformer **S-HBg**− is characterized by an internal H-bond (OH···O= distance: 1.695 Å, Fig. 4). In this arrangement the NH_2 group becomes pseudo axial with respect to the $OC_1C_2C_3OH$ cycle and cannot approach the other carboxylate oxygen at a distance lower than 2.382 Å (Fig. 4). The net result is a conformer slightly destabilized by comparison with **S-H$_A$a** and **S-H$_B$a**. At the G4 level the ratio of the conformer populations **S-H$_B$a**/**S-H$_A$a**/**S-H$_B$g**−/**S-H$_A$g**+ is equal to 52.0/43.2/4.5/0.3.

In previous theoretical studies devoted to deprotonated serine,[28,31] conformers **S-H$_A$a** and **S-H$_B$a** have been identified as the most stable on the basis of B3LYP computations. Recently, experimental gas phase infrared multiphoton dissociation (IRMPD) spectra of $[Ser - H]^-$ ions have been recorded in the 600–1800 cm^{-1} range.[36] The authors interpret their results by using frequencies calculated at the B3LYP/6-31++G(d,p) level (and scaled by 0.98) and concluded that conformers **S-H$_A$a** and **S-H$_B$a** are the most likely structures of deprotonated serine in the gas phase in perfect agreement with our data.

Threonine. The conformational landscape of neutral L-threonine has been the subject of recent theoretical investigations.[25,37–39] A number of conformers situated between 56[38] and 74[39] have been found using the B3LYP functional and either 6-311++G(d,p) or 6-31G(d) basis sets. Our investigation of (2*S*,3*R*)-L-threonine leads to the characterization of a similar range of 65 stable structures. Full data concerning the twelve most stable conformers are given in the ESI (Table T2 and Fig. S7 and S8).† As already observed for serine, composite and M06-2X methods predict comparable relative H°_{298} and G°_{298} energies. By contrast, B97-D and B3LYP functionals may lead to H°_{298} or G°_{298} with a variance of several kJ mol^{-1} (Table T2, Fig. S7).† The difference observed between the relative stability orders, when passing from H°_{298} to G°_{298}, is also easily understood when considering the structural differences between conformers. Since at the G4 level five conformers (**TII$_B$g**−, **TI$_C$g**+, **TI$_A$g**−, **TIII$_C$g**+, **TI′$_C$g**−) are predicted to represent

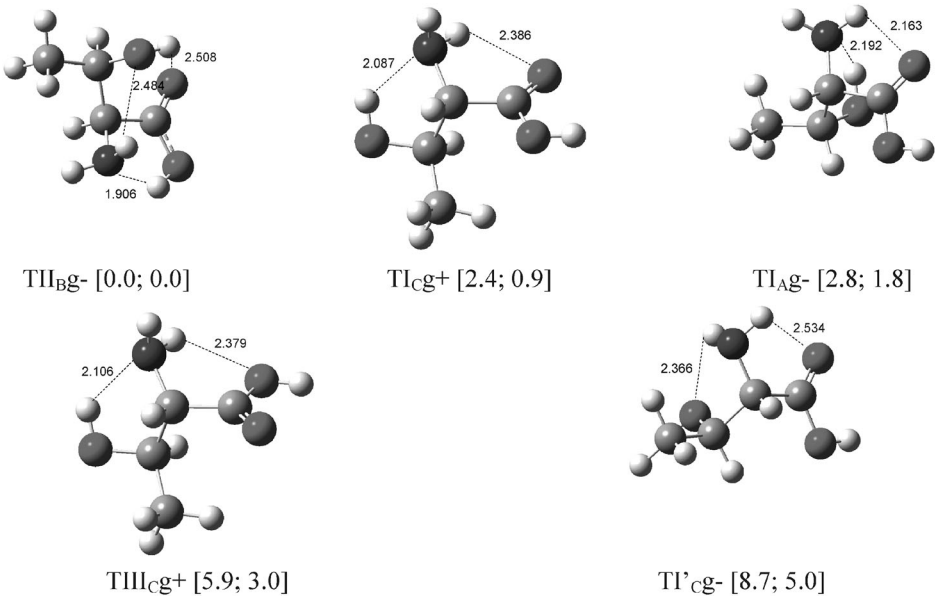


Fig. 6 Optimized geometries (B3LYP/6-31G(2df,p)) of the five most stable conformers of neutral L-threonine (in brackets, relative H°_{298} and G°_{298} in kJ mol^{-1} calculated at the G4 level).

~90% of the population of neutral threonine at 298 K, only these structures are presented in Fig. 6 and discussed below.

As observed for serine, conformers of type **II** present the strongest internal hydrogen bonding and consequently the lowest enthalpy but also the lowest absolute entropies. For serine and threonine, this situation arises for conformers **SII$_B$g**− and **TII$_B$g**−. It is noteworthy that the same order in enthalpy is observed for the first conformers of both serine and threonine (*i.e.* **II$_B$g**− < **I$_C$g**+ < **I$_A$g**− < **II$_B$a**). However, for serine the enthalpy gap between **SII$_B$g**− and **SI$_C$g**+ is equal to 1.2 kJ mol^{-1} while for threonine, the analogous difference attains 2.4 kJ mol^{-1} (G4 calculations). Since the entropies of type **II** conformers are lower than the entropies of conformers of type **I**, the G°_{298} of both types of conformers will be affected differently. Roughly, relative G°_{298} are shifted downward by *ca.* 1.8 kJ mol^{-1}. As a consequence, the order of stability is changed for **SII$_B$g**− and **SI$_C$g**+ but not for **TII$_B$g**− and **TI$_C$g**+.

As indicated above, conformers **TII$_B$g**−/**TI$_C$g**+/**TI$_A$g**−/**TIII$_C$g**+/**TI'$_C$g**− represent more than 90% of the population of conformers at 298 K (G4 calculations). The exact ratio of **TII$_B$g**−/**TI$_C$g**+/**TI$_A$g**−/**TIII$_C$g**+/**TI'$_C$g**− is 38.0/26.7/18.8/11.4/5.1%. Comparable results, based on MP2/6-311+G(2df,p) calculations, have been reported by Lin and coworkers.[39] From an experimental point of view, Fourier transform microwave spectroscopy was applied to L-threonine vaporized by laser ablation.[40] Analysis of the resulting rotational spectra leads the authors to conclude that seven conformers were present in their experimental conditions, the most abundant of which being **TI$_C$g**+, **TII$_B$g**− and **TI$_A$g**− in excellent agreement with our expectations.

Relative enthalpies and free energies at 298 K of the four most stable structures of protonated threonine, among the 12 identified stable structures at the B3LYP/6-31+G(d,p) level, are reported in Table T2 and Fig. S9 and S10 of ESI.† Not surprisingly, as for serine, essentially two conformers, **THIg**− and **THIg**+, are predicted to represent the equilibrium population at 298 K. These two structures present indeed particularly strong intramolecular hydrogen bonds between the NH_3^+ moiety and (i) the oxygen of the carbonyl group, and (ii) the hydroxyl oxygen (Fig. 7). Two other structures, **THIIg**+ and **THIIg**−, characterized by a less efficient $NH_3^+\cdots$O(H)–C=O bonding are situated more than 10 kJ mol^{-1} above **THIg**− and **THIg**+. This situation is reminiscent of that encountered with serine and similar comments can be made. In particular the small compression of the energy scale observed between H°_{298} and G°_{298} (Fig. S9 of ESI†) is in perfect agreement with the difference in strength of the internal hydrogen bonds and the resulting difference in entropy. A comparable situation has been described before with protonated serine (see Fig. 3 and accompanying comments). Finally, it is evident that the description of the 298 K population of protonated threonine is limited to the participation of conformers **THIg**− and **THIg**+ only. Accordingly, using the G4 computed G°_{298}, a **THIg**−/**THIg**+/**THIIg**+/**THIIg**− ratio of 61.2/38.1/0.4/0.3% is evaluated.

Concerning deprotonated threonine, we located 11 stable conformers, the four most stable being similar to those obtained for serine, **T-H$_A$a**, **T-H$_B$a**, **T-H$_B$g**−, and **T-H$_A$g**+ (Table T2 and Fig. S11 and S12 of ESI).† Their relative H°_{298} and G°_{298} (Fig. 7) present similar values and the choice of the method seemingly does not alter the corresponding stability orders. A noticeable exception however is the B97-D functional which doesn't predict **T-H$_A$a** as a stable species but invariably converges toward conformer **T-H$_B$a**. The structural difference between **T-H$_A$a** and **T-H$_B$a** is only lying on the value of the HNC=O dihedral angle (Fig. 7). The slightly lower stability of **T-H$_B$a** is probably due to a steric repulsion between the methyl group and the H atom of the amino group not involved in a hydrogen bond. This argument is supported by the fact that, in serine, the two equivalent conformers **S-H$_A$a** and **S-H$_B$a** present quasi identical H°_{298} and G°_{298} (Fig. 4).

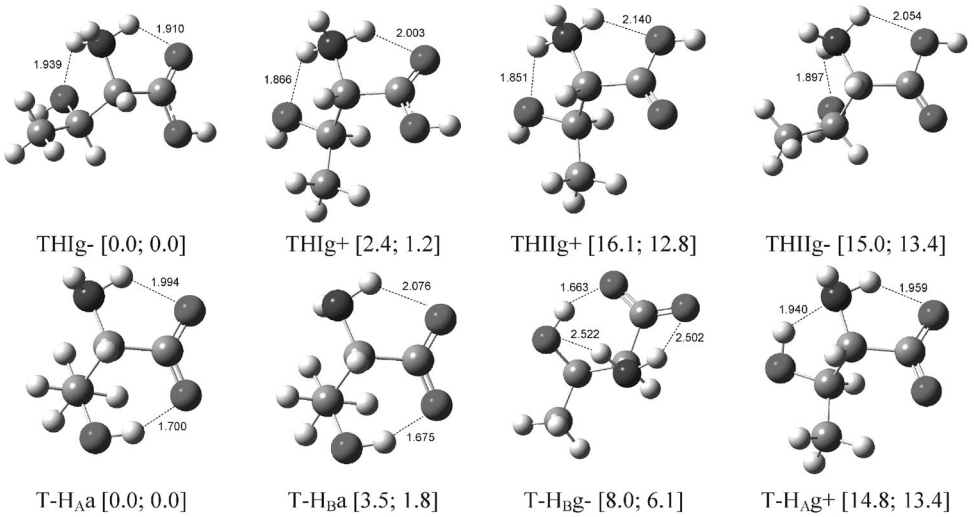


Fig. 7 Optimized geometries (B3LYP/6-31G(2df,p)) of the four most stable conformers of protonated and deprotonated L-threonine (in brackets, relative H°_{298} and G°_{298} in kJ mol^{-1} calculated at the G4 level).

Using the G4 free energies at 298 K, we estimate the population of deprotonated threonine, **T-H$_A$a**/**T-H$_B$a**/**T-H$_B$g**−/**T-H$_A$g**+, to 62.9/31.2/5.6/0.3%.

Cysteine. Among the 87 stable conformers of (2*R*)-L-cysteine identified at the DFT level, the 13 most stable have been investigated with the composite computational methods. The relative H°_{298} and G°_{298} and the optimized structures of these conformers are presented in Table T3 and Fig. S13 and S14 of ESI.† As a general observation for this system, relative 298 K enthalpies of the ten most stable conformers are close together whatever the computational method used. More significant fluctuations are however observed on the relative free energies. However, all the methods give the type **II** conformer **CII$_B$g**− (Fig. 8) as the most stable in both H°_{298} and G°_{298}.

At the G4 level, the order of stability of the cysteine conformers in terms of H°_{298} is: **CII$_B$g**− < **CI′$_C$g**− < **CI$_C$g**− < **CI$_A$g**−, **CI$_C$g**+ < **CII$_A$g**+ < **CIII$_C$g**− (Fig. 8). It is noteworthy that the enthalpy difference between **CII$_B$g**− and the second conformer **CI′$_C$g**− is as large as 5.4 kJ mol^{-1} at the G4 level. No such large energy gap has been observed for serine and threonine for which conformer **II$_B$g**− has been also identified as the most stable in the enthalpy scale, but where the closest conformers were situated only 1.2 and 2.4 kJ mol^{-1} above. This difference in behaviour may be understood by examining the network of internal hydrogen bonds stabilizing the concerned conformers. As repeatedly noted, conformers of type **II** are strongly stabilized by the H$_2$N···HOC=O(*anti*) H bonding. A secondary stabilization is brought by a NH···OHC(3) interaction for **SII$_B$g**− and **TII$_B$g**−. In the case of **CII$_B$g**−, the homologous interaction, *i.e.* NH···SHC(3) is reinforced since S is more basic than O (proton affinities of methanol and methanethiol are 754 and 773 kJ mol^{-1}, respectively). Conformers of type **I** are characterized by a NH···O=COH(*syn*) interaction. In addition, **I$_A$g**− and **I$_C$g**+ conformers of serine and threonine enjoy very strong C(3)OH···NH$_2$ interactions. This results in close enthalpy proximity for conformers **SII$_B$g**−, **SI$_A$g**− and **SI$_C$g**+ on one hand and of **TII$_B$g**−, **TI$_A$g**− and **TI$_C$g**+ on the other. In the case of cysteine, the stabilization of conformers **CI$_A$g**− and **CI$_C$g**+ is not observed because the C(3)SH···NH$_2$ interaction becomes less efficient due to the lower electronegativity of S.

The difference in stability between **CII$_B$g**− and the other conformers is significantly reduced in the free energy scale. The order of stability given by the G4 calculated G°_{298}: **CII$_B$g**− < **CI$_C$g**− < **CI′$_C$g**− < **CI$_C$g**+ < **CI$_A$g**− < **CIII$_C$g**− < **CII$_A$g**+ is however not significantly changed with respect to the enthalpy scale. What is noteworthy is the shift of the conformers of type **I** (**CI$_C$g**−, **CI′$_C$g**−, **CI$_C$g**+, **CI$_A$g**−) toward low G°_{298} as expected from their high third law entropies (near 380 J mol^{-1} K^{-1}, as compared to 367 J mol^{-1} K^{-1} for **CII$_B$g**−). This entropy difference is obviously related to the strong character of the two major internal hydrogen bonds occurring in **CII$_B$g**− discussed above. The six conformers presented in Fig. 8 represent *ca.* 85% of the total population of conformers (Table T3 in ESI). At 298 K the overall ratio **CII$_B$g**−/**CI$_C$g**−/**CI′$_C$g**−/**CI$_C$g**+/**CI$_A$g**−/**CIII$_C$g**− is predicted to be 35.8/23.9/15.1/11.9/7.4/5.9%.

Previous theoretical works on neutral cysteine reported results obtained at the HF,[24] B3LYP[26,29,41–43] and MP2[44,45] levels of theory. Some of these studies however are concerned by the non-natural 2*S* configuration of cysteine.[24,26,41,45] There is a general consensus in placing conformer **CII$_B$g**− as the most stable species. Depending upon the theoretical level, the conformer situated immediately above **CII$_B$g**− is **CI$_C$g**+[24,29,41,43] or **CI$_C$g**−[44,45] as expected for a structure differing only by few kJ mol^{-1}. Experimentally, no less than six conformers have been claimed to be identified.[44] Accordingly, microwave spectra of vapor phase cysteine have been interpreted by the major presence of conformers **CII$_B$g**−, **CI$_C$g**+ and **CI$_C$g**− beside the minor contributions of conformers **CII$_A$g**+, **CIII$_C$a** and **CIII$_C$g**−.[44] Surprisingly enough, conformers **CI′$_C$g**− and **CI$_A$g**− are not mentioned despite the fact that they are predicted to be more stable than **CII$_A$g**+, **CIII$_C$a** or **CIII$_C$g**−. It may be underlined however that rotational constants of **CI′$_C$g**− and **CI$_A$g**− are very close to that of **CI$_C$g**− probably rendering a clear structural assignment uneasy.

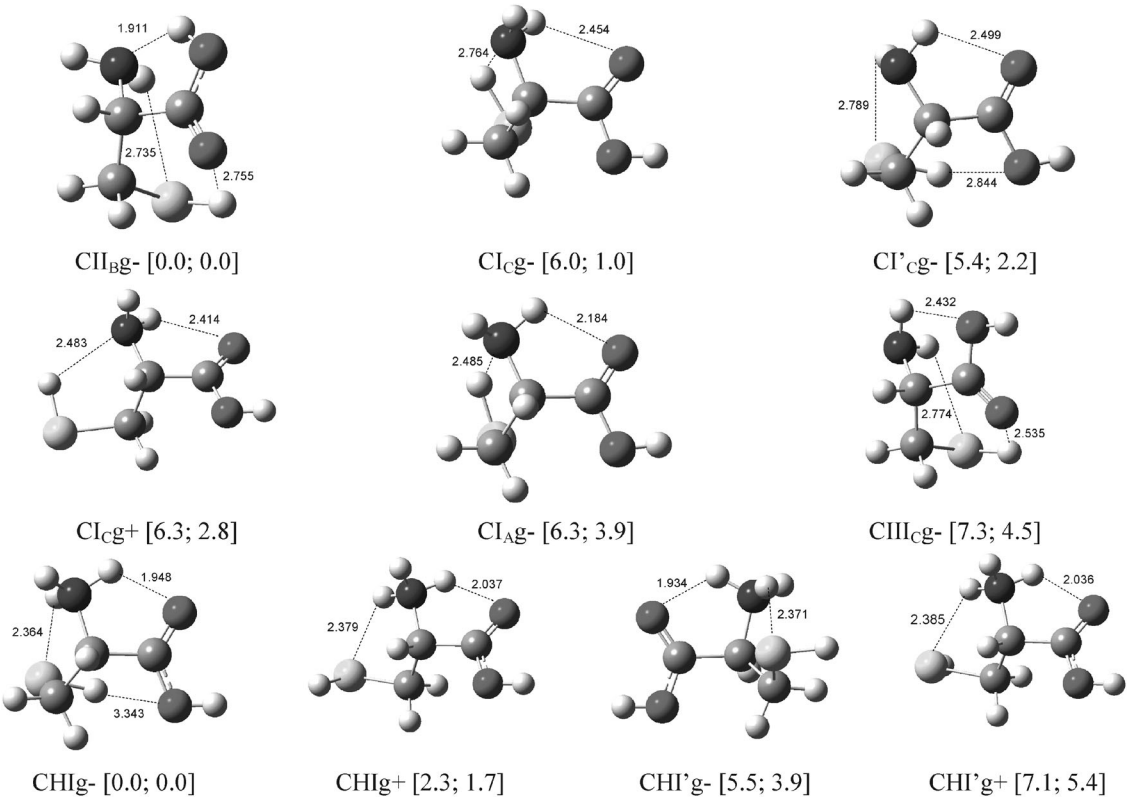


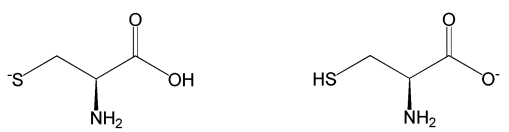
Fig. 8 Optimized geometries (B3LYP/6-31G(2df,p)) of the six stable conformer of neutral L-cysteine and the four most stable conformers protonated L-cysteine (in brackets, relative H°_{298} and G°_{298} in kJ mol^{-1} calculated at the G4 level).

Among the 21 conformers identified for protonated cysteine, four were predicted to be situated in a ~6 kJ mol^{-1} range in both H°_{298} and G°_{298} (Table T3 and Fig. S15, S16 in ESI,† Fig. 8). These four structures, **CHIg**−, **CHIg**+, **CHI′g**− and **CHI′g**+, are stabilized by $NH_3^+\cdots$O=COH(*syn*) and $NH_3^+\cdots$S hydrogen bonding interactions. The difference between conformers **CHIg**− and **CHI′g**−, or **CHIg**+ and **CHI′g**+, lies on a rotation around the C(3)–S bond. For **CHIg**− and **CHIg**+ the S–H bond is staggered with respect to the C(3)H$_2$,C(2) group (Fig. 8), whereas the conformation is eclipsed for conformers **CHI′g**− and **CHI′g**+. This conformational change is associated with a difference in energy of *ca.* 5 kJ mol^{-1}. The **g**−/**g**+ conformational change corresponds to an energy difference of ~2 kJ mol^{-1} seemingly induced by a slight weakening of the $NH_3^+\cdots$O=COH(*syn*) interaction in **CHIg**+ and **CHI′g**+. The four conformers described above represent the essential components of the equilibrium population at 298 K. The distribution of conformers **CHIg**−/**CHIg**+/**CHI′g**−/**CHI′g**+, computed using the G4 free energies, is equal to 54.2/28.1/11.4/6.3%. The present results are in qualitative agreement with a previous study, done, as indicated previously, on the 2*R* enantiomeric form of cysteine.[26]

Deprotonated cysteine may exist in either its carboxylate or its thiolate forms (Scheme 2). Recently, mass spectrometry[43] and photoelectron spectroscopy[46] experiments were interpreted by the preferential formation of the thiolate form of deprotonated cysteine. By contrast, tentative characterization of the thiolate structure by gas phase IR multi photon dissociation failed but, rather, the results were interpreted by the occurrence of a carboxylate form.[36] These studies were supported by theoretical computations on several deprotonated cysteine conformers where the thiolate structure is lower in energy than the carboxylate structure. However only a limited number of conformers seems to have been considered in these studies. Moreover, most of the time, the investigations were done using the B3LYP functional. A systematic investigation of the conformational space of deprotonated cysteine up to the G4 level is thus of interest in order to elucidate these observations.

In the present study we identified 25 carboxylate, **C-H$_O$**, and 16 thiolate, **C-H$_S$**, stable structures, the five most stable are situated in a 15 kJ mol^{-1} G°_{298} range (Table T3 and Fig. S17, S18 in ESI,† Fig. 9). All the composite methods (CBS-QB3, G3B3, G4MP2 and G4) confirm that the thiolate form is more stable than the carboxylate form of deprotonated cysteine.

In fact, four thiolate conformers, **C-H$_{SC}$g**−, **C-H$_{SC}$a**, **C-H$_{SA}$a**, **C-H$_{SB}$g**−, are predicted to be situated below the lowest



C-H$_S$ **C-H$_O$**

Scheme 2

carboxylate form **C-H$_{OC}$g+** (Fig. 9). The four thiolate structures are all characterized by a strong (*anti*)OCOH$\cdots^{-}$S– interaction. Secondary favourable interactions involving $NH_2\cdots O{=}COH$ or $NH_2\cdots{}^{-}S$– hydrogen bonding are observed for the *a* (**C-H$_{SC}$a** and **C-H$_{SA}$a**) or *g*– (**C-H$_{SC}$g**– and **C-H$_{SB}$g**–) types of conformers, respectively. These four structures are situated in a 6 kJ mol^{-1} range of H°_{298} and appear to be similarly stabilized by their networks of internal hydrogen bonds since they present identical third law entropies of 355 J K^{-1} mol^{-1}. As a consequence, the relative G°_{298} values mimic the H°_{298} results.

The carboxylate structure **C-H$_{OC}$g+** is stabilized by a $NH_2\cdots{}^{-}O{=}CO$ internal hydrogen bond, accompanied by a cooperative $SH\cdots NH_2$ interaction. This structure is situated 12.8 kJ mol^{-1} above the most stable thiolate conformer **C-H$_{SC}$g**– in the H°_{298} scale. This enthalpy gap is drastically reduced when considering the G°_{298}. Accordingly, since the entropy of **C-H$_{OC}$g+** is 14 J K^{-1} mol^{-1} larger than that of **C-H$_{SC}$g**–, the difference in free energy is reduced to 8.7 kJ mol^{-1} at the G4 level. It is confirmed however that the carboxylate form is not the most stable, even on the 298 K free energy scale. The predicted population of conformers **C-H$_{SC}$g**–/**C-H$_{SC}$a**/**C-H$_{SA}$a**/**C-H$_{SB}$g**–/**C-H$_{OC}$g+** is 57.2/28.7/8.1/4.2/1.8% (G4 calculations at 298 K). It is noteworthy that only **C-H$_{SC}$a** and **C-H$_{OC}$g+** are considered in ref. 36 and 43. It would be consequently interesting to re-examine the photoelectron and IRMPD results in light of the present results which demonstrate the presence of conformers **C-H$_{SC}$g**–, **C-H$_{SA}$a** or **C-H$_{SB}$g**–.

Methionine. The neutral methionine potential energy surface has been examined by several groups at various levels of theory.[30–33,48–51] Most of the time, a limited number of conformers has been considered. The most recent investigation reports results obtained at the B3LYP, B3P86 and MP2 levels using a 6-311+G(2d,2p) basis set and conclude that six conformers are lying in the first 5 kJ mol^{-1} of the H°_{0} range.[51] In the present study, the 32 most stable conformers obtained at the B3LYP/6-31+G(d,p) level were further investigated with the whole panel of methods used here (Table T4 and Fig. S19 and S20 in ESI).† Fig. 10 presents the six most stable conformations in the G°_{298} scale identified in the present work at the G4 level. These structures are unambiguously of type **I**: **MetI$_{C}$g+g+g+**, **MetI$_{C}$g+ag–**, **MetI$_{C}$g+g+a**, **MetI$_{C}$g+ag+**, **MetI$_{C}$g+aa** and **MetI$_{C}$g+g+g–** (Fig. 10). These structures are also among the most stable in the H°_{298} scale but three type **II** conformers fall in the same 5 kJ mol^{-1} energy range: **MetII$_{B}$g–g+g+**, **MetII$_{B}$g–g+a** and **MetII$_{A}$g+g–g–** (Fig. S20).† All the computational methods agree in locating **MetI$_{C}$g+g+g+** as the most stable conformer (Table T4 in ESI).† It is remarkable that this finding has been reported only in one previous study.[33] The stability of structures **MetI$_{C}$g+g+g+**, **MetI$_{C}$g+ag–**, **MetI$_{C}$g+g+a**, **MetI$_{C}$g+ag+**, **MetI$_{C}$g+aa** and **MetI$_{C}$g+g+g–** is due to the $NH_2\cdots O{=}C$ interaction (type **I** conformers) and also to a $C(4)H\cdots NH_2$ favourable interaction, attested by a small interatomic distance ($d \sim 2.6$ Å), and due to the occurrence of a positive charge of *ca.* +0.2 on the hydrogen atom and a negative charge of *ca.* −0.7 on the nitrogen atom. The largest stability of **MetI$_{C}$g+g+g+** may be attributed to an additional interaction between the C(2)H hydrogen (charge +0.22) and the sulfur atom (charge −0.22) their separating distance being only 2.9 Å while it attains more than 4 Å for the other **MetI$_{C}$g+** conformers. Conformers of type **II**, **MetII$_{B}$g–g+a**, **MetII$_{A}$g+g–g–** and **MetII$_{B}$g–g+g+**, present a NH$\cdots$S hydrogen bonding type interaction (interatomic distance ~2.5 Å) which significantly reduces the backbone flexibility and thus the absolute entropy. The S°_{298} values of **MetII$_{B}$g–g+a**, **MetII$_{A}$g+g–g–** and **MetII$_{B}$g–g+g+** are indeed close to 435 J mol^{-1} K^{-1} *i.e.* 15 J mol^{-1} K^{-1} less than the S°_{298} values of type **I** conformers thus explaining why these latter are shifted

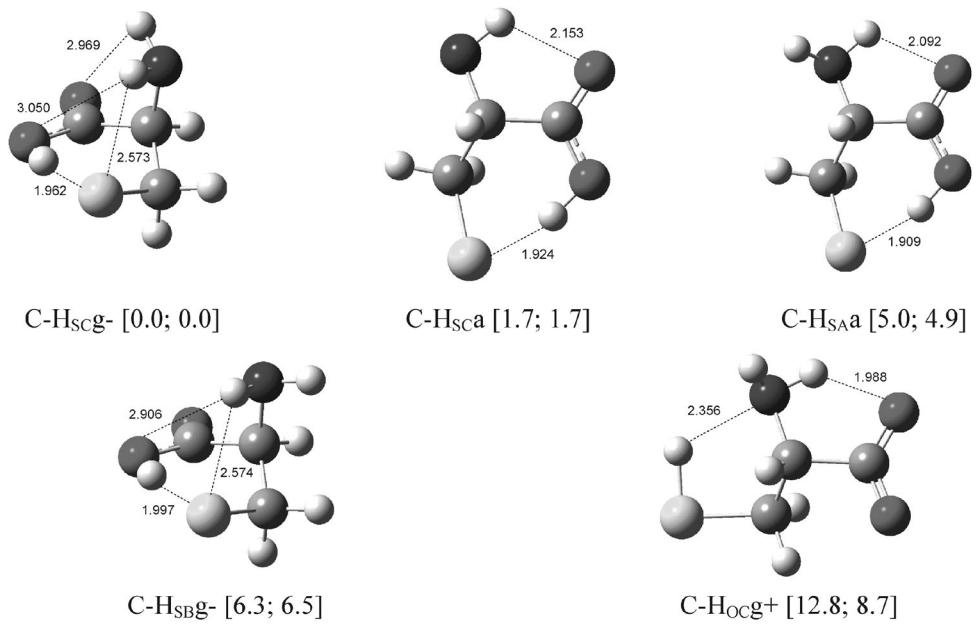
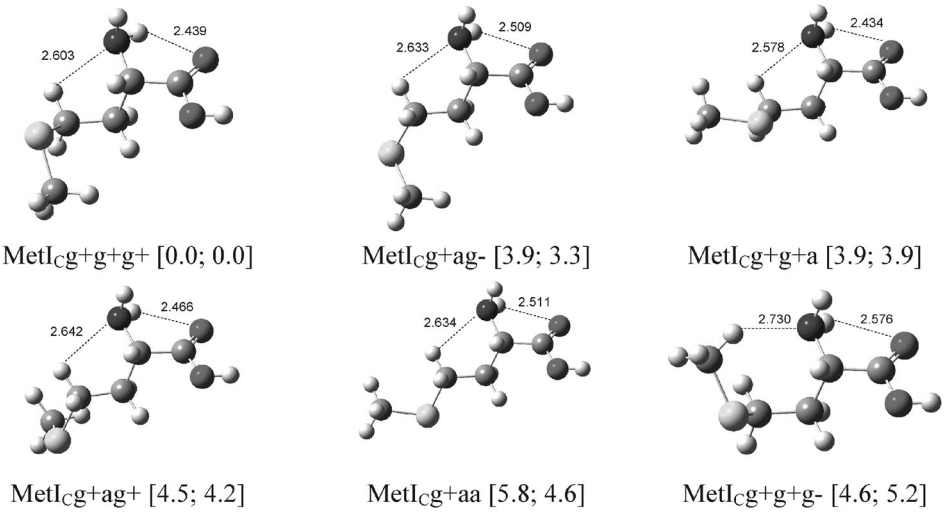


Fig. 9 Optimized geometries (B3LYP/6-31G(2df,p)) of the five most stable conformers of deprotonated L-cysteine (in brackets, relative H°_{298} and G°_{298} in kJ mol^{-1} calculated at the G4 level).



MetI$_C$g+g+g+ [0.0; 0.0] MetI$_C$g+ag- [3.9; 3.3] MetI$_C$g+g+a [3.9; 3.9]

MetI$_C$g+ag+ [4.5; 4.2] MetI$_C$g+aa [5.8; 4.6] MetI$_C$g+g+g- [4.6; 5.2]

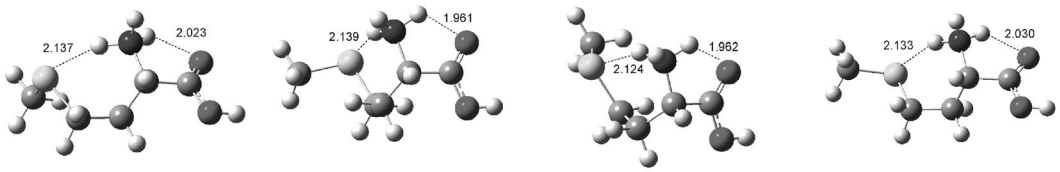
Fig. 10 Optimized geometries (B3LYP/6-31G(2df,p)) of the six most stable conformers of neutral L-methionine (in brackets, relative H°_{298} and G°_{298} in kJ mol^{-1} calculated at the G4 level).

to the low G°_{298} values. The theoretical population of neutral methionine conformers at 298 K calculated at the G4 level is **MetI$_C$g+g+g+**/**MetI$_C$g+ag−**/**MetI$_C$g+g+a**/**MetI$_C$g+ag+**/**MetI$_C$g+aa**/**MetI$_C$g+g+g−**: 51.2/13.8/10.8/9.4/8.3/6.5%.

Neutral methionine has been studied experimentally by valence core photoelectron spectroscopy in the VUV and soft X-ray regions[48] and by gas phase Fourier transform IR spectroscopy.[47] However, observations were interpreted by computations on conformers situated more than 10–15 kJ mol^{-1} above the most stable form **MetI$_C$g+g+g+**, re-examination of the data in the light of the present results is consequently suggested.

Only two conformers of protonated methionine were considered previously.[30,32,33,49,50] The present study reveals that four conformers of type **HI**: **MetHIg+g−g−**, **MetHIg−g+g+**, **MetHIg−g+a** and **MetHIg+g−a** are situated in a ~3 kJ mol^{-1} H°_{298} and G°_{298} range. Conformers of type **HII**, namely **MetHIIg+g−g−** and **MetHIIg+g−a** are situated more than 10 kJ mol^{-1} above the four above mentioned homologues of type **HI** (Table T4 and Fig. S21, S22 in ESI).† Beside the $NH_3^+\cdots$OCOH(*syn*) (type **HI**) and $NH_3^+\cdots$OHCO(*syn*) (type **HII**) interactions, the six structures are stabilized by a second internal hydrogen bond involving the sulfur atom as a proton acceptor: $NH_3^+\cdots$S. It is noteworthy that the NH$\cdots$S distance is remarkably constant for the six conformers (~2.134 ± 0.007 Å) thus suggesting a comparable stabilization effect. The large energy difference observed between conformers of type **HI** and **HII** (~10 kJ mol^{-1}) is in line with identical observations done for Ser, Thr and Cys and is obviously due to the difference in basicity of both oxygens of the acidic group. Clearly, conformers **MetHIg+g−g−** and **MetHIg−g+g+** are of identical stabilities and should be hardly distinguishable (Fig. 11). At the G4 level, the calculated relative population of conformers at 298 K, **MetHIg+g−g−**/**MetHIg−g+g+**/**MetHIg−g+a**/**MetHIg+g−a** is equal to 33.8/29.8/22.1/14.3%.

The 24 most stable conformers of deprotonated methionine were investigated with the total panoply of computational methods (Table T4 and Fig. S23 and S24 of ESI).† At the G4 level, six structures were found to fall in the 0–5 kJ mol^{-1} G°_{298} range (**Met-H$_A$g−ag+**, **Met-H$_B$g−ag+**, **Met-H$_A$g+ag+**, **Met-H$_A$g+ag−**, **Met-H$_A$g−ag−**, **Met-H$_A$g−aa**), two additional structures (**Met-H$_A$ag−g+**, **Met-H$_B$ag−g+**) were found to be situated in the 0–5 kJ mol^{-1} range of the H°_{298} scale (Fig. 12). All the investigated conformers present the HNH$\cdots^-$O=CO interaction with an interatomic distance of ~2.0 Å. It is generally observed that this distance is shorter for conformers of type **Met-H$_A$** which, as a consequence, are more stable than conformers of type **Met-H$_B$** (by 1 to 5 kJ mol^{-1}) (Fig. 12). Additional stabilizations are provided by interactions of the types (i) C(4)H$\cdots^-$O=CO (**Met-H$_A$g−ag+**, **Met-H$_B$g−ag+**, **Met-H$_A$g−ag−**, **Met-H$_A$g−aa**, **Met-H$_A$ag−g+**, **Met-H$_B$ag−g+**), (ii) C(4)H$\cdots$NH$_2$ (**Met-H$_A$g−ag+**, **Met-H$_A$g+ag+**, **Met-H$_A$g+ag−**, **Met-H$_A$g−ag−**, **Met-H$_A$g−aa**)



MetHIg+g-g- [0.0; 0.0] MetHIg-g+g+ [0.5; 0.3] MetHIg-g+a [2.0; 1.1] MetHIg+g-a [3.2; 2.2]

Fig. 11 Optimized geometries (B3LYP/6-31G(2df,p)) of the four most stable conformers of protonated L-methionine (in brackets, relative H°_{298} and G°_{298} in kJ mol^{-1} calculated at the G4 level).

and (iii) $SCH_3\cdots{}^{-}O{=}{=}CO$ (**Met-H$_B$g−ag+**, **Met-H$_A$ag−g+**, **Met-H$_B$ag−g+**). It is noteworthy that conformers **Met-H$_A$ag−g+** and **Met-H$_B$ag−g+** are characterized by the shortest $H\cdots{}^{-}O{=}{=}CO$ distance (~2.3 Å and ~2.1 Å for interactions of types (i) and (iii), respectively) thus explaining their position at the lower part of the enthalpy scale. These strong internal hydrogen bonds have also the consequence of reducing the absolute entropies of these conformers. Indeed, the S°_{298} values of **Met-H$_A$ag−g+** and **Met-H$_B$ag−g+** are equal to 426 J mol^{-1} K^{-1} while all the other conformers present S°_{298} values close to 440 J mol^{-1} K^{-1}. For this reason, the two conformers are shifted by *ca.* 5 kJ mol^{-1} upward in the Gibbs free energy scale. Using the G4 G°_{298} values, the population of conformers **Met-H$_A$g−ag+/Met-H$_B$g−ag+/Met-H$_A$g+ag+/Met-H$_A$g+ag−/Met-H$_A$g−ag−/Met-H$_A$g−aa** is predicted to be 47.1/18.0/10.8/10.1/8.1/5.9%.

3.2 Protonation and deprotonation thermochemistry

Experimental proton affinities and gas phase basicities of Ser, Thr, Cys and Met are summarized in Table 1 while theoretical values are reported in Table 2. Similar information concerning $\Delta_{acid}H$ and $\Delta_{acid}G$ are presented in Tables 3 and 4. It may be underlined that computed PA and GB values reported in Tables 2 and 4 are defined with respect to the most stables conformers in the H°_{298} and G°_{298} scales, respectively. It may consequently correspond to different conformers, particularly for the neutral species (these particular cases are marked with an asterisk in Tables 2 and 4).

Experimental. The general principle of measurement of gas-phase protonation, and deprotonation, thermochemistry is to determine the Gibbs free energy change associated with a proton transfer between the molecule M of interest and a reference base B_{ref}. Three general methods have been developed, (i) the equilibrium method, (ii) the kinetic methods ("simple" and "extended") and (iii) the thermokinetic method.[1–3]

The first determinations of the gas phase basicities of Ser, Thr, Cys and Met came from the measurement of proton transfer equilibrium constants by ion cyclotron resonance mass spectrometry.[52] In these experiments, the amino acid molecules M have been volatilized by heating the sample in a direct insertion probe close to the reacting region. No measurement of the temperature has been done but an estimate of 350 K has been proposed by Hunter & Lias.[53] The derived GB values quoted in Table 1 take into account this temperature correction. Serine and methionine were also examined by the thermokinetic[54] and the extended kinetic[49,50] methods. These procedures lead to GB values in good agreement with those obtained by the equilibrium method. Average GB values are indeed associated with reasonable standard deviation (less than 3 kJ mol^{-1}).

The "simple" kinetic method is expected to provide an apparent proton affinity value given by $PA_{app}(M) = [PA(M) - T_{eff}\{\Delta_pS^{\circ}(M) - \Delta_pS^{\circ}(B_{ref})\}]$ where T_{eff} is an "effective temperature" and $\Delta_pS^{\circ}(X) = S^{\circ}(XH^{+}) - S^{\circ}(X)$ the "protonation entropy" of the species X (M or B_{ref}).[3,53] In the "simple" kinetic method it is assumed that the experimentally determined

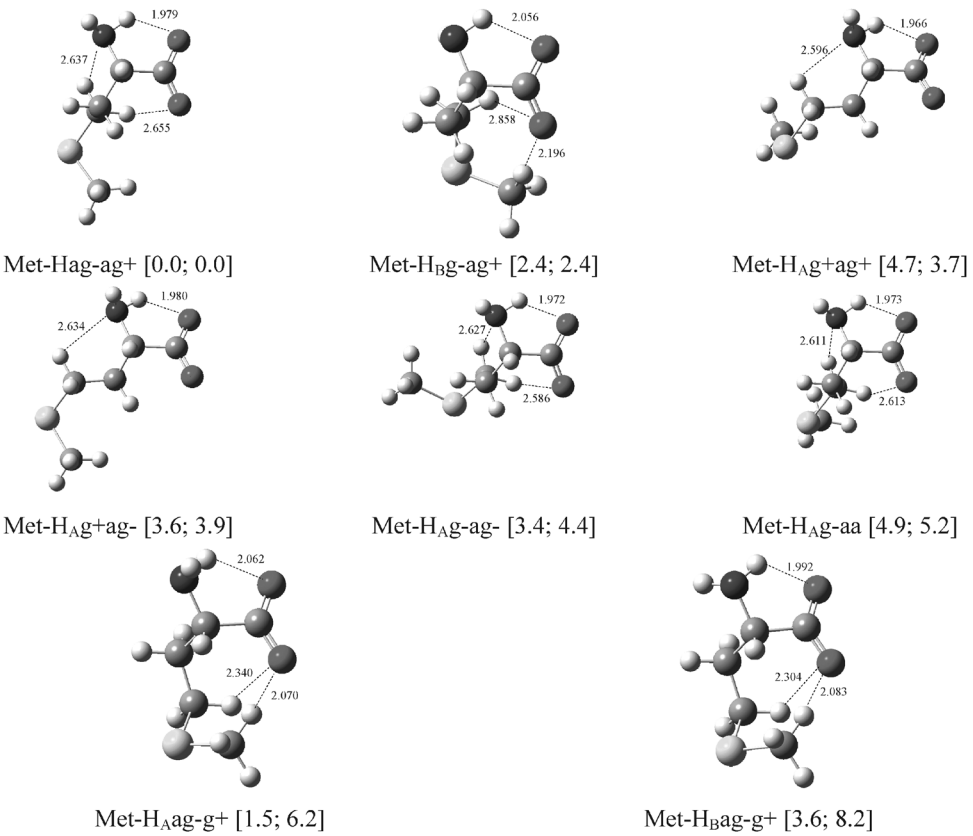


Fig. 12 Optimized geometries (B3LYP/6-31G(2df,p)) of the six most stable conformers + two conformers of deprotonated L-methionine (in brackets, relative H°_{298} and G°_{298} in kJ mol^{-1} calculated at the G4 level).

Table 1 Experimental protonation thermochemistry of the studied amino acids[a]

M	Method	GB(M) (kJ mol^{-1})	PA(M) (kJ mol^{-1})	$\Delta_pS°$(M)[b] (J K^{-1} mol^{-1})
Serine	Equilibrium	880.3[c]		
	Thermokinetic	876.2 ± 4.3[d]		
	Simple kinetic		910.4[e]	
			912.9[f]	
			911.4[g]	
	Average	878.3 ± 2.9	911.6 ± 1.3	
	Evaluated	880.7[j]–874.3[k]	914.6[j]–906.7[k]	−5[j]–0[k]
Threonine	Equilibrium	888.5[c]		
	Simple kinetic		918.7[e]	
			922.8[f]	
			921.4[g]	
	Average	888.5	921.0 ± 2.1	
	Evaluated	888.5[j]–893.5[k]	922.5[j]–925.9[k]	−5[j]–0[k]
Cysteine	Equilibrium	868.8[c]		
	Simple kinetic		904.0[e]	
			901.9[f]	
	Average	868.8	903.0 ± 1.5	
	Evaluated	869.3[j]–868.9[k]	903.2[j]–901.4[k]	−5[j]–0[k]
Methionine	Equilibrium	901.3[c]		
	Bracketing	899.0		
	Simple kinetic		928.4[e]	
			931.6[f]	
			927.2[h]	
			936.5[g]	
			928.7[i]	
			(930.5 ± 3.7)	
	Extended kinetic	898.5 ± 3.2[i]	937.5 ± 2.9[i]	−22 ± 5[i]
			924.7–931.4[h]	
	Average	899.6 ± 1.5	931.7 ± 4.4	
	Evaluated	901.5[j]–900.6[k]	935.4[j]–933.0[k]	−5[j]–0[k]

[a] Data anchored to GB(NH_3) = 819.0 kJ mol^{-1} and PA(NH_3) = 853.6 kJ mol^{-1}, and corrected to the Hunter & Lias[53] basicity scale using linear correlation with the original data. [b] $\Delta_pS°(M) = S°(MH^+) - S°(M)$. [c] Ref. 52 (original values and, in brackets, as adapted by Hunter & Lias[53] by assuming a temperature of 350 K rather than 320 K). [d] Ref. 54. [e] Ref. 57. [f] Ref. 58. [g] Ref. 59. [h] Ref. 49. [i] Ref. 50. [j] Evaluated by Hunter & Lias.[53] [k] Ref. 60.

$PA_{app}(M)$ is equal to the true PA(M). It consequently supposes that the term $T_{eff}\{\Delta_pS°(M) - \Delta_pS°(B_{ref})\}$ can be neglected. In situations where this quantity is not negligible a more elaborated method such as the "extended kinetic method" is needed. The latter allows the determination of both parameters PA(M) and $\{\Delta_pS°(M) - \Delta_pS°(B_{ref})\}$ by considering several sets of experiments corresponding to different effective temperatures T_{eff} *i.e.* to different ion activation conditions.[3] This latter method has been applied only to methionine and revealed a negative $\Delta_pS°$(Met) of *ca.* −20 J K^{-1} mol^{-1}. It may be noted that the PA_{app} values obtained by the "simple" kinetic method for methionine from five different laboratories (Table 1) are ranging from 928 to 936 kJ mol^{-1} and present a standard deviation of 4.4 kJ mol^{-1}. By comparison, the standard deviations associated to the PA_{app} values of Ser, Thr and Cys are situated between 1.3 and 2.1 kJ mol^{-1}. The spread of PA_{app}(Met) values observed in Table 1 may be interpreted by different effective temperatures T_{eff} associated with the various experiments. The existence of a non negligible $\Delta_pS°$(Met) term would consequently lead to different PA_{app} values. By the same reasoning one would expect negligible $\Delta_pS°$ values for the three amino acids Ser, Thr and Cys.

Gas phase acidities, $\Delta_{acid}G$, of serine, threonine, cysteine and methionine have been determined by using the kinetic methods by O'Hair *et al.*[55] and by Poutsma and coworkers.[31] They anchor their results to the $\Delta_{acid}G$ of reference molecules including substituted benzoic acids. In the most recent compilation of thermochemical data concerning negative ions,[56] the $\Delta_{acid}G$ values of the reference acids used by O'Hair *et al.*[55] have changed. We thus reconsider their original experimental data and adjust the resulting $\Delta_{acid}G$ to the new gas phase acidity scale.[56] $\Delta_{acid}H$ reported in Table 3 are obtained either from the sets of experimental $\Delta_{acid}G$ obtained by the simple kinetic method[31,55] assuming a constant $T\Delta S$ term of 29.5 kJ mol^{-1}, or from data derived from the extended kinetic method.[31] Comparison of the various $\Delta_{acid}G$ and $\Delta_{acid}H$ estimates (Table 3) shows a clear agreement from one method to the other with an average deviation equal to 2.7 kJ mol^{-1}.

Theoretical. Monoconformer proton affinities and gas phase basicities, PA_{msc} and GB_{msc}, of Ser, Thr, Cys and Met computed using the most stable conformers in the enthalpy scale and in the free energy scale, respectively are compared in Table 2. Consideration of the overall population of conformers at 298 K for both neutral and protonated species leads to averaged proton affinities, $PA_{average}$, protonation entropies, $\Delta_pS°_{average}$, and gas phase basicities, $GB_{average}$. The two latter quantities include the entropy of mixing (eqn (11)) in the "average" values. It should be underlined that "average" values of the thermochemical parameters are dependent on the number of conformers and consequently on the Gibbs free energy range considered. We adopt uniformly a cutoff of ΔG = 6 kJ mol^{-1} since, using eqn (7) at 298 K, any conformer more than 6 kJ mol^{-1} will possess a relative abundance less than 10% that of the most stable conformer.

It is generally observed that PA_{msc} constitutes a lower limit for the overall, averaged, quantities $PA_{average}$. This is due to the larger number of conformers for the neutral species with respect to the protonated form in a given energy range. It results in a difference $[H°_{298}(M)]_{average} - H°_{298}(M)_{msc}$ larger than $[H°_{298}(MH^+)]_{average} - H°_{298}(MH^+)_{msc}$ and consequently to $PA_{average} > PA_{msc}$. The difference $PA_{average} - PA_{msc}$ observed in the present study is equal to 2.2 ± 0.6 kJ mol^{-1} at the G4 level (Table 2). When looking at the gas phase basicity, the data presented in Table 2 show that $GB_{average}$ is lower than GB_{msc} by 1.5 kJ mol^{-1} for Ser, Thr and Cys at the G4 level. It is noteworthy that the $\Delta_pS°_{average}$ of these three amino acids reduces to the entropy of mixing contribution: the intrinsic $\Delta_pS°$ differences calculated either in the monoconformer approximation or by considering the full population of conformers are, on average, close to zero. In the case of methionine the situation is different since a $\Delta_pS°$ of *ca.* −20 J K^{-1} mol^{-1} is calculated (both using the msc or average methods). As a consequence, $GB_{average}$ is higher than GB_{msc} (by 1.4 kJ mol^{-1} at the G4 level).

It is now interesting to compare the thermochemical results obtained at the various levels of theory investigated here using the G4 method as a benchmark. Signed average deviations (SAD) on GB(M) and PA(M) are reported in Table T5 of the ESI.† The largest average deviations are obtained using B97-D and M06-2X functionals. Accordingly, PA_{msc} calculated at the B97-D/6-311++G(3df,2p) level are always overestimated,

Table 2 Computed protonation thermochemistry of the studied amino acids[a]

		GB(M) (kJ mol^{-1})		PA(M) (kJ mol^{-1})		ΔS°(M)[b] (J K^{-1} mol^{-1})	
M	Method	msc	Average	msc	Average	msc	Average
Serine	B97-D/6-31+G(d,p)	888.8*(876.2)	888.1(875.5)	919.0*(906.5)	920.2(907.7)	2.4	1.3
	B97-D/6-311++G(3df,2p)	887.7(882.0)		917.8(912.2)		8.2	
	B3LYP/6-31+G(d,p)	879.2(876.3)	879.0(876.1)	913.0(910.2)	913.0(910.2)	−4.4	−4.7
	B3LYP/6-311++G(3df,2p)	878.2*(882.9)		911.9*(916.7)		−4.4	
	M06-2X/6-31+G(d,p)	873.6*(873.6)	872.9(872.9)	906.4*(906.6)	906.7(906.9)	−3.9	−4.4
	M06-2X/6-311++G(3df,2p)	869.6*(879.9)		901.7*(912.1)		−3.9	
	CBS-QB3	878.2*(877.4)	876.4(875.6)	910.3*(909.6)	911.3(910.6)	1.3	−8.2
	G3B3	880.1*(877.8)	878.6(876.3)	912.4*(910.3)	913.5(911.4)	−4.5	−8.2
	G4MP2	880*(879.1)	878.9(878.0)	913.1*(912.4)	913.8(913.1)	−5.1	−8.3
	G4	**879.9*(878.3)**	**878.6(877.0)**	**912.7*(911.2)**	**913.6(912.1)**	**−5.1**	−8.2
	Experiment average	878.3 ± 2.9		911.6 ± 1.3			
Threonine	B97-D/6-31+G(d,p)	896.1*(883.5)	895.9(883.3)	925.6*(913.1)	927.2(914.7)	10.6	3.8
	B97-D/6-311++G(3df,2p)	895.2*(889.5)		925.1*(919.5)		10.6	
	B3LYP/6-31+G(d,p)	888.2(885.3)	887.7(884.8)	920.1(917.3)	920.9(918.1)	1.9	−2.5
	B3LYP/6-311++G(3df,2p)	886.6(891.3)		918.5(923.3)		1.9	
	M06-2X/6-31+G(d,p)	881.7(881.7)	880.4(880.4)	911.9(912.1)	913.3(913.5)	7.6	−1.4
	M06-2X/6-311++G(3df,2p)	878.2(888.5)		908.5(918.9)		7.6	
	CBS-QB3	883.9(883.1)	883.5(882.7)	915.7(915.0)	916.6(915.9)	2.3	−2.3
	G3B3	887.2(884.9)	885.9(883.6)	918.1(916.0)	919.8(917.7)	5.3	−4.9
	G4MP2	887.3(886.4)	885.9(885.0)	919.0(918.3)	920.5(919.8)	2.7	−6.9
	G4	**886.9(885.3)**	**885.7(884.1)**	**918.6(917.1)**	**919.9(918.4)**	**2.7**	−5.9
	Experiment average	888.5		921.0 ± 2.1			
Cysteine	B97-D/6-31+G(d,p)	883.1(870.5)	883.4(870.8)	913.3(900.8)	915.3(902.8)	7.7	2.2
	B97-D/6-311++G(3df,2p)	884.1(878.4)		914.3(908.7)		7.7	
	B3LYP/6-31+G(d,p)	871.5(868.6)	870.7(867.8)	903.3(900.5)	906.7(903.9)	2.2	−11.7
	B3LYP/6-311++G(3df,2p)	872.6*(877.3)		904.3*(909.1)		5.1	
	M06-2X/6-31+G(d,p)	862.6(862.6)	861.8(861.8)	893.9(894.1)	896.4(896.6)	4.1	−6.9
	M06-2X/6-311++G(3df,2p)	861.8*(872.1)		893.0*(903.4)		11.3	
	CBS-QB3	870.2(869.4)	868.8(868.0)	901.8(901.1)	904.3(903.6)	3.0	−10.2
	G3B3	872.2(869.9)	870.7(868.4)	903.2(901.1)	905.7(903.6)	4.8	−8.4
	G4MP2	872.9(872.0)	871.4(870.5)	904.4(903.7)	907.1(906.4)	3.2	−10.6
	G4	**872.1(870.5)**	**870.9(869.3)**	**903.6(902.1)**	**906.1(904.6)**	**3.2**	−9.0
	Experiment average	868.8		903.0 ± 1.5			
Methionine	B97-D/6-31+G(d,p)	923.4*(910.8)	921.5(908.9)	957.8*(945.3)	960.0(947.5)	−10.7	−20.1
	B97-D/6-311++G(3df,2p)	923.1*(917.4)		957.9*(952.3)		−10.7	
	B3LYP/6-31+G(d,p)	905.3(902.4)	904.1(901.2)	942.8(940.0)	943.9(941.1)	−16.7	−24.5
	B3LYP/6-311++G(3df,2p)	904.8(909.5)		942.3(947.1)		−16.7	
	M06-2X/6-31+G(d,p)	895.9*(895.9)	895.0(895.0)	931.0*(931.2)	933.0(933.2)	−2.9	−18.7
	M06-2X/6-311++G(3df,2p)	894.2(904.5)		929.2(939.6)		−8.4	
	CBS-QB3	898.5(897.7)	899.7(898.9)	936.5(935.8)	937.7(937.0)	−18.5	−18.3
	G3B3	901.3(899.0)	902.6(900.3)	939.1(937.0)	940.4(938.3)	−17.7	−17.5
	G4MP2	900.4(899.5)	901.3(900.4)	938.8(938.1)	939.7(939.0)	−19.5	−19.9
	G4	**900.3(898.7)**	**901.4(899.8)**	**938.7(937.2)**	**939.7(938.2)**	**−19.5**	−19.7
	Experiment average	899.6 ± 1.5		937.5 ± 2.9		−22 ± 5	

[a] "most stable conformer" (msc) value and averaged (average) values calculated over the 298 K distribution of conformers (entropy of mixing is included in the "average" GB(M) and $\Delta_p S^{\circ}$(M) values). An asterisk means that the most stable conformer in enthalpy is different from the most stable in Gibbs free energy at 298 K. Values in parentheses are anchored to the tabulated PA(NH_3) = 853.6 kJ mol^{-1} and GB(NH_3) = 819.0 kJ mol^{-1}.[53] [b] $\Delta_p S^{\circ}(M) = S^{\circ}(MH^+) - S^{\circ}(M)$.

mainly for the sulfur containing molecules (average deviation: 10.4 kJ mol^{-1}, with a maximum deviation equal to 19 kJ mol^{-1} for methionine), while M06-2X/6-311++G(3df,2p) PA_{msc} are systematically underestimated by ~10 kJ mol^{-1}. Interestingly enough, B3LYP/6-311++G(3df,2p) gives the correct PA_{msc} (within less than 2 kJ mol^{-1}). Excellent agreement is found between G3B3, G4MP2 and G4 with a maximum SAD of 0.4 kJ mol^{-1}. Note that CBS-QB3 leads to PA_{msc} values slightly below that given by the G4 method (by −2.5 kJ mol^{-1}). Similar conclusion arises from examination of the monoconformer gas phase basicities, GB_{msc}.

Computed monoconformers and averaged $\Delta_{acid}G$, $\Delta_p S^{\circ}$ and $\Delta_{acid}H$ are presented in Table 4. As done for the protonation thermochemistry, a cutoff of $\Delta G^{\circ}_{298} = 6$ kJ mol^{-1} has been applied to limit the number of conformers considered in the 298 K population averaging. Averaged $\Delta_{acid}H_{average}$ appears to be slightly lower than monoconformers $\Delta_{acid}H_{msc}$ by *ca.* 2 kJ mol^{-1} (Table 4). As noted for proton affinities, this shift finds its origin in the larger number of conformers of the neutral molecules with respect to their ionized forms in the 298 K populations. The resulting increase in enthalpy of one mol of conformational mixture of neutral amino acid consequently reduces $\Delta_{acid}H$. Turning now to $\Delta_{acid}G$, the data presented in Table 4 show that $\Delta G_{acid,\ average}$ is slightly higher than $\Delta_{acid}G_{msc}$, the shift is however less than 1 kJ mol^{-1}.

Taking the G4 results as benchmark, the maximum signed average deviation SAD on $\Delta_{acid}H$ is obtained with the B97-D/6-311++G(3df,2p) data ($\Delta_{acid}H_{msc}$ ~ +7 kJ mol^{-1}) (Table T6 of ESI).† A systematic underestimate of $\Delta_{acid}H_{msc}$

Table 3 Experimental deprotonation thermochemistry of aliphatic α-amino acids[a]

M	Method	$\Delta_{acid}G$(M) (kJ mol^{-1})	$\Delta_{acid}H$(M) (kJ mol^{-1})
Serine	Simple kinetic[b]	1364.5	1394.0
	Simple kinetic[c]	1365.0	1394.3
	Extended kinetic[d]	1360.9	1391
	Average	1363.5 ± 2.2	1393.1 ± 1.8
Threonine	Simple kinetic[b]	1361.8	1391.3
	Simple kinetic[c]	1363.8	1393.4
	Extended kinetic[d]	1360.2	1388
	Average	1361.9 ± 1.8	1390.9 ± 2.7
Cysteine	Equilibrium[e]	1370.3 ± 8.8	1399.1 ± 9.2
	Simple kinetic[b]	1365.5	1395.0
	Simple kinetic[c]	1369.8	1399.3
	Extended kinetic	1366.5[d]	1395[c]
		1364.6(±14)[e]	1392.9(±14)[e]
	Average	1367.3 ± 2.6	1396.3 ± 2.8
Methionine	Simple kinetic[b]	1379.4	1408.9
	Simple kinetic[c]	1384.8	1414.3
	Extended kinetic[d]	1378.1	1407
	Average	1380.7 ± 3.6	1410.1 ± 3.8

[a] Data anchored to $\Delta_{acid}G$(benzoic acid) = 1394.0 kJ mol^{-1} and $\Delta_{acid}H$(benzoic acid) = 1423.5 kJ mol^{-1} (ref. 56). [b] Ref. 55, $\Delta_{acid}G$ recalculated using the most recent $\Delta_{acid}G$ of the reference acids, $\Delta_{acid}H$(M) were obtained by adding a uniform $T\Delta S$ term equal to 29.5 kJ mol^{-1}. [c] Using the data of ref. 31 obtained at 20% attenuation. [d] $\Delta_{acid}H$(M) from ref. 31, $\Delta_{acid}G$(M) are calculated using the Δ_pS(M − H) = S^o(M) − S^o([M − H]$^-$) = 8 ± 35, 16 ± 20, 13 ± 19 and 12 ± 9 J mol^{-1} K^{-1} for Ser, Thr, Cys and Met, respectively (personal communication from JC Poutsma). [e] Ref. 43.

(SAD ∼ −4 kJ mol^{-1}, Table S22) is obtained using B3LYP/6-311++G(3df,2p), M06-2X/6-311++G(3df,2p) or CBS-QB3 results. A close agreement is found between G3B3 and G4 results (SAD = −0.8 kJ mol^{-1}) while G4MP2 provides $\Delta_{acid}H_{msc}$ slightly higher than that given by the G4 method (SAD = +1.8 kJ mol^{-1}). Comparable results are obtained for the monoconformer gas phase acidities, $\Delta_{acid}G_{msc}$ (Table T6 of ESI).†

Comparison experiment/theory. Computed thermochemical data reported in Tables 2 and 4, either in the "msc" or the "average" approximations, are ideal quantities which should be compared carefully with experiment. It is a custom to compare data obtained by mass spectrometry techniques to computed "msc" values. However, the population of ions sampled experimentally does not necessarily consist of a pure single conformer or even to a thermal distribution of structures. It depends on how the ions are formed and handled during the experiments. In equilibrium or kinetic methods, proton exchange is occurring into proton bonded complexes stabilized by *ca.* 50–100 kJ mol^{-1}. This deep energy well allows, in principle, the interconversion of conformers through rotational barriers along σ-bonds and consequently all thermodynamically accessible isomers to be present. Recently, evidence was presented on the validity of the "average" approach (including entropy of mixing) when considering proton transfer equilibrium experiments.[61] No such evidence has been reported for data obtained by using the kinetic method. A second point to consider is the ability of a given theoretical method to reproduce the thermochemical parameters of the standards to which the protonation/deprotonation thermochemistry is anchored. As indicated in Tables 1 and 3, the reference standards are ammonia for the basicity scale and benzoic acid for the acidity scale. Computed PA(NH_3), GB(NH_3), $\Delta_{acid}H$(benzoic acid) and $\Delta_{acid}G$(benzoic acid) using the panel of methods used in the present work are reported in Table T9 of ESI† together with the presently recognized reference values.[53,56] Deviations larger than 12 kJ mol^{-1} are observed on PA(NH_3) and GB(NH_3) whereas maximum deviations less than 6 kJ mol^{-1} are obtained for $\Delta_{acid}H$(benzoic acid) and $\Delta_{acid}G$(benzoic acid). Correction based on these deviations are applied to the computed PA, GB, $\Delta_{acid}H$ and $\Delta_{acid}G$ values and are indicated in parentheses in the data reported in Tables 2 and 4.

When comparing experimental and G4 computed proton affinities reported in Table 2,a correct agreement appears using both monoconformer PA_{msc} or averaged $PA_{average}$ values (SAD_{msc} = −1.4 ± 1.7 kJ mol^{-1} against $SAD_{average}$ = 0.1 ± 1.6 kJ mol^{-1}). A clear illustration is given by Fig. 13 where G4 calculated proton affinities are plotted against the experimental values. As evidenced in Fig. 13, the maximum deviation is observed for threonine, suggesting that the experimental value is too high by *ca.* 3 kJ mol^{-1}. In the case of methionine, only the experimental PA value obtained by the extended kinetic method is meaningful since a significant protonation entropy is demonstrated for this molecule. Indeed, the calculated G4 values (PA_{msc} = 937.2 and $PA_{average}$ = 938.2 kJ mol^{-1}) agree nicely with experiment (PA_{exp} = 937.5 ± 2.9 kJ mol^{-1}). Comparison of gas phase basicities is also correct when considering G4 calculated GB_{msc} or $GB_{average}$ and experimental data (SAD_{msc} = −0.6 ± 2.0 kJ mol^{-1} against $SAD_{average}$ = −1.3 ± 2.2 kJ mol^{-1}). Again, the maximum deviation is observed for threonine and one may suggest that the experimental GB(threonine) is too high by *ca.* 3 kJ mol^{-1}. Comparison between experimental and computed protonation entropy can be done only for methionine. It is observed (Table 2) that the Δ_pS^o value obtained by the extended kinetic method (−22 ± 5 J mol^{-1} K^{-1})[50] is in excellent agreement with predictions based on the computations either using the msc or the average results (−20 J mol^{-1} K^{-1} Table 2). No experimental information is available on the protonation entropy of serine, threonine and cysteine but, as recalled above, it is generally assumed negligible.[53] In agreement with this expectation, computed Δ_pS^o are generally limited to less than ±10 J mol^{-1} K^{-1}: the mean msc and average values are 0.3 ± 4.7 and −7.7 ± 1.6 J mol^{-1} K^{-1}, respectively.

Experimental $\Delta_{acid}H$ of Ser, Thr, Cys and Met are generally slightly closer to the computed monoconformer $\Delta_{acid}H_{msc}$ values (SAD_{msc} = 0.5 ± 2.0 and $SAD_{average}$ = −0.8 ± 3.2 kJ mol^{-1}) (Table 4 and Table T7 of ESI).† Similar conclusions may be drawn when considering $\Delta_{acid}G$ (SAD_{msc} = 0.5 ± 1.6 and $SAD_{average}$ = 1.2 ± 0.9 kJ mol^{-1}). In both cases, the maximum deviation is observed for methionine indicating that the experimental $\Delta_{acid}H$ and $\Delta_{acid}G$ values are probably too low by ∼3 kJ mol^{-1}. Considering protonation entropy Δ_pS^o([M − H]$^-$) associated with gas phase acidity, values situated between 8 and 16 J mol^{-1} K^{-1} were obtained from extended kinetic method plots,[31] the corresponding 95% error was however of the same order of magnitude (see footnote *c*

Table 4 Computed deprotonation thermochemistry of aliphatic α-amino acids[a]

		$\Delta_{acid}G$(M) (kJ mol^{-1})		$\Delta_{acid}H$(M) (kJ mol^{-1})		Δ_pS^o(M − H)[b] (J K^{-1} mol^{-1})	
M	Method	msc	Average	msc	Average	msc	Average
Serine	B97-D/6-31+G(d,p)	1360.4*(1361.1)	1361.5(1362.2)	1391.3*(1390.7)	1390.7(1390.1)	10.6	11.2
	B97-D/6-311++G(3df,2p)	1366.3(1363.8)		1397.4(1393.6)		4.8	
	B3LYP/6-31+G(d,p)	1352.3(1358.8)	1352.4(1358.9)	1381.0(1386.7)	1380.2(1385.9)	13.0	15.8
	B3LYP/6-311++G(3df,2p)	1356.9*(1360.4)		1385.6*(1388.3)		13.0	
	M06-2X/6-31+G(d,p)	1355.6*(1361.4)	1356.4(1362.2)	1385.8*(1390.6)	1384.6(1389.4)	11.0	14.4
	M06-2X/6-311++G(3df,2p)	1358.5*(1363.5)		1389.2*(1393.2)		11.0	
	CBS-QB3	1356.5*(1360.4)	1357.7(1361.6)	1386.8*(1390.2)	1384.9(1388.3)	12.8	17.8
	G3B3	1360.4*(1360.8)	1361.3(1361.7)	1390.5*(1390.4)	1388.4(1388.3)	13.1	18.2
	G4MP2	1364.0*(1364.6)	1364.4(1365.0)	1393.0*(1393.2)	1391.7(1391.9)	14.8	17.5
	G4	**1361.8*(1363.3)**	**1362.4(1363.9)**	**1391.1*(1392.2)**	**1389.7(1390.8)**	**14.8**	17.3
	Experiment average	1363.5 ± 2.2		1393.1 ± 1.8		8 ± 35	
Threonine	B97-D/6-31+G(d,p)	1363.2(1363.9)	1364.4(1365.1)	1393.6(1393.0)	1392.7(1392.1)	7.1	14.0
	B97-D/6-311++G(3df,2p)	1368.8(1366.3)		1399.2(1395.4)		7.1	
	B3LYP/6-31+G(d,p)	1353.0(1359.5)	1354.0(1360.5)	1383.2(1388.9)	1383.2(1388.9)	7.9	11.1
	B3LYP/6-311++G(3df,2p)	1356.8(1360.3)		1386.9(1389.6)		7.9	
	M06-2X/6-31+G(d,p)	1354.9*(1360.7)	1355.6(1361.4)	1386.4*(1391.2)	1385.1(1389.9)	−2.1	10.2
	M06-2X/6-311++G(3df,2p)	1357.6(1362.6)		1389.1(1393.1)		3.4	
	CBS-QB3	1355.8(1359.7)	1356.8(1360.7)	1386.1(1389.5)	1385.0(1388.4)	7.5	14.4
	G3B3	1358.7(1359.1)	1360.7(1361.1)	1389.5(1389.4)	1387.6(1387.5)	5.6	18.9
	G4MP2	1362.3(1362.9)	1363.7(1364.3)	1392.2(1392.4)	1391.4(1391.6)	8.6	16.1
	G4	**1360.4(1361.9)**	**1361.6(1363.1)**	**1390.3(1391.4)**	**1389.6(1390.7)**	**8.6**	15.2
	Experiment average	1361.9 ± 1.8		1390.9 ± 2.7		16 ± 20	
Cysteine	B97-D/6-31+G(d,p)	1365.6(1366.3)	1364.9(1365.6)	1394.3(1393.7)	1393.0(1392.4)	12.7	23.1
	B97-D/6-311++G(3df,2p)	1373.3(1370.8)		1402.0(1398.2)		12.7	
	B3LYP/6-31+G(d,p)	1354.7(1361.2)	1356.0(1362.5)	1384.3(1390)	1382.0(1387.7)	9.5	21.8
	B3LYP/6-311++G(3df,2p)	1363.0(1366.5)		1392.7(1395.4)		9.5	
	M06-2X/6-31+G(d,p)	1348.9(1354.7)	1349.7(1355.5)	1379.2(1384.0)	1375.8(1380.6)	7.5	21.6
	M06-2X/6-311++G(3df,2p)	1356.1(1361.1)		1386.4(1390.4)		7.5	
	CBS-QB3	1362.0(1365.9)	1362.6(1366.5)	1391.2(1394.6)	1388.3(1391.7)	11.1	23.0
	G3B3	1363.4(1363.8)	1364.6(1365.0)	1392.9(1392.8)	1389.5(1389.4)	10.1	25.6
	G4MP2	1366.9(1367.5)	1368.6(1369.2)	1396.0(1396.2)	1392.8(1393.0)	11.5	27.7
	G4	**1365.2(1366.7)**	**1366.7(1368.2)**	**1394.3(1395.4)**	**1391.1(1392.2)**	**11.5**	27.2
	Experiment average	1367.3 ± 2.6		1396.3 ± 2.8		13 ± 19	
Methionine	B97-D/6-31+G(d,p)	1379.3*(1380.0)	1379.9(1380.6)	1411.8*(1411.2)	1411.8(1411.2)	−2.5	2.2
	B97-D/6-311++G(3df,2p)	1383.6*(1381.1)		1416.4*(1412.6)		−2.5	
	B3LYP/6-31+G(d,p)	1371.9*(1378.4)	1371.8(1378.3)	1405.1*(1410.8)	1404.2(1409.9)	−2.3	0.1
	B3LYP/6-311++G(3df,2p)	1374.2(1377.7)		1407.4(1410.1)		−2.3	
	M06-2X/6-31+G(d,p)	1374.8*(1380.6)	1374.3(1380.1)	1405.5*(1410.3)	1406.6(1411.4)	−17.3	0.9
	M06-2X/6-311++G(3df,2p)	1377.6*(1382.6)		1408.5*(1412.5)		−17.3	
	CBS-QB3	1377.8(1381.7)	1377.4(1381.3)	1408.8(1412.2)	1408.1(1411.5)	5.2	6.1
	G3B3	1381.3(1381.7)	1381.1(1381.5)	1411.9(1411.8)	1411.5(1411.4)	6.4	7.0
	G4MP2	1383.8(1384.4)	1383.8(1384.4)	1414.1(1414.3)	1413.9(1414.1)	7.5	6.2
	G4	**1382.1(1383.6)**	**1381.6(1383.1)**	**1412.3(1413.4)**	**1412.3(1413.4)**	**7.5**	6.3
	Experiment average	1380.7 ± 3.6		1410.1 ± 3.8		12 ± 9	

[a] "most stable conformer" (msc) value and averaged (average) values calculated over the 298 K distribution of conformers (entropy of mixing is included in the "average" $\Delta_{acid}G$(M) and Δ_pS^o(M − H) values). An asterisk means that the most stable conformer in enthalpy is different from the most stable in Gibbs free energy at 298 K. Values in parentheses are anchored to the tabulated ΔacidG(benzoic acid) = 1394.0 kJ mol^{-1} and $\Delta_{acid}H$(benzoic acid) = 1423.5 kJ mol^{-1}.[56] [b] Δ_pS^o(M − H) = S^o(M) − S^o(M − H).

in Table 3). The computed $\Delta_pS^o([M - H]^-)$ presented in Table 4 are in correct agreement with these experimental data and confirm the existence of a small, but systematic, positive $\Delta_pS^o([M - H]^-)$ associated with the deprotonation of Ser, Thr, Cys and Met. This conclusion may be compared to the observation of a $\Delta_pS^o([M - H]^-)$ close to 10 J mol^{-1} K^{-1} for alkyl substituted amino acids.[23]

4. Conclusion

The present work was supported by an extensive search of the most stable conformers of neutral, protonated and deprotonated serine, threonine, cysteine and methionine. From a large investigation of more than around two thousand trial geometries based on systematic dihedral angle changes completed by Monte-Carlo and molecular dynamics simulations procedures using AMOEBA force field, no less than 131 structures were fully examined at various theoretical levels. These explorations involved three DFT (B3LYP, B97-D and M06-2X functionals using 6-31+G(d,p) optimized geometries and 6-311++G(3df,2p) single points) and four composite (CBS-QB3, G3B3, G4MP2 and G4) methods.

Comparison between these various computational methods leads to the following observations:

(a) Conformers' relative energies calculated by the composite methods CBS-QB3, G3B3, G4MP2, G4 and by the M06-2X functional are generally nearly identical (within ~2 kJ mol^{-1}). Larger discrepancies are observed when using B97-D or B3LYP functionals, particularly for the neutral species.

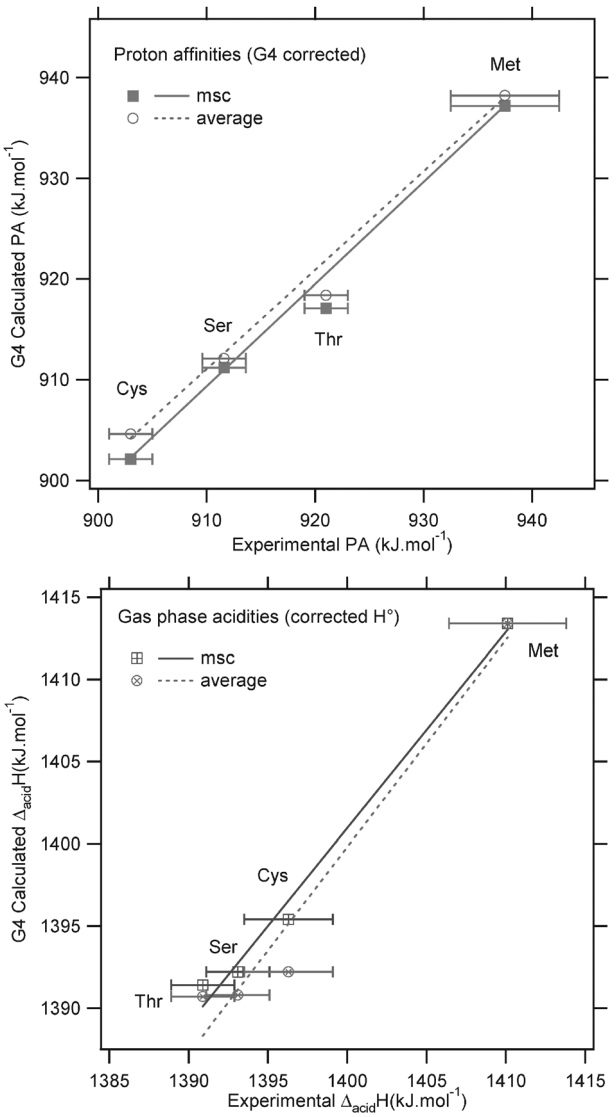


Fig. 13 Examples of linear correlations observed between G4 computed and experimental thermochemical data (a): proton affinities and (b): enthalpy acidities.

(b) Monoconformer proton affinities computed by the G3B3, G4MP2 and G4 are practically identical (standard deviation 0.4 kJ mol^{-1}). B3LYP/6-311++G(3df,2p) PA_{msc} are remarkably close to the G4 values but the two other functionals B97-D and M06-2X do not work so nicely since an average deviation of $\sim$10 kJ mol^{-1} is observed.

(c) Monoconformer $\Delta_{acid}H$ computed by the composite G4 and G3B3 methods are identical, systematic shifts of +1.8 kJ mol^{-1} and −2.5 kJ mol^{-1} are observed with G4MP2 and CBS-QB3 methods, respectively. DFT methods are less accurate in computing $\Delta_{acid}H$ by $\sim$7 kJ mol^{-1} for B97-D, $\sim$4 kJ mol^{-1} for B3LYP and $\sim$2 kJ mol^{-1} for M06-2X (with respect to the G4 values).

When comparison between theory and experiment is done, several important findings may be underlined:

(a) Experimental and G4 computed PA and $\Delta_{acid}H$ compare satisfactorily, a slightly better agreement is observed if the computed monoconformer "msc" values are considered. The deviation with average PA and $\Delta_{acid}H$ is however limited to 1.5 kJ mol^{-1}.

(b) Computed Δ_pS^o(M) are generally limited to less than $\pm$10 J mol^{-1} K^{-1} except for methionine for which a value close to −20 J mol^{-1} K^{-1} is calculated (both in the msc and the average estimates) in agreement with experiment.

(c) Assignment of the structure(s) (or mixtures of structures) of the most stable neutral conformer(s) is in correct agreement with experiments based on microwave spectra as attested for serine,[27b] threonine[40] and cysteine.[44] By contrast, interpretation of the experimental IR[47] and VUV photoelectron[48] spectra of neutral methionine should be reconsidered by taking into account the structure of the most stable conformers identified here. A similar remark applies to the IRMPD spectra of protonated serine and deprotonated cysteine. In the former case, only one structure (namely **SHIg**−) has been considered to interpret the spectrum;[35] our results show that a second conformer should also participate (namely **SHIg**+). Its impact on the observed absorption bands is however expected to be negligible since both **SHIg**− and **SHIg**+ exhibit the same characteristic frequencies in the considered spectral zone. Concerning deprotonated cysteine, our computational results confirm that the thiolate form is the most stable.[36,43,46] The IRMPD spectrum of [cysteine − H]$^-$[36] should be however reinterpreted by considering the structures evidenced here as the most stable *i.e.* mainly **C-HSCg**−.

Finally, based on the comparison between experiment and G4 theory, a set of evaluated thermochemical data for serine, threonine, cysteine, and methionine may be proposed: PA = 912, 918, 903, 938; GB = 879, 886, 870, 900; $\Delta_{acid}H$ = 1392, 1391, 1395, 1413; $\Delta_{acid}G$ = 1363, 1362, 1367, 1383 kJ mol^{-1}.

Acknowledgements

The authors would like to thank Bhawani Singh Inda for preliminary calculations, Ashwani Sharma and Carine Clavaguéra for their help during the setup of the computational procedures. This work was performed using HPC resources from GENCI-CINES (grant 2011-c2011085107).

References

1 K. M. Ervin, *Chem. Rev.*, 2001, **101**, 391–444.
2 J. F. Gal, P. C. Maria and E. D. Raczynska, *J. Mass Spectrom.*, 2001, **36**, 699–716.
3 G. Bouchoux, *Mass Spectrom. Rev.*, 2007, **26**, 775–835.
4 M. Alcami, O. Mo and M. Yanez, *Mass Spectrom. Rev.*, 2001, **20**, 195–245.
5 M. Alcami, O. Mo and M. Yanez, *J. Phys. Org. Chem.*, 2002, **15**, 174–186.
6 (*a*) K. Müller-Dethlefs and P. Hobza, *Chem. Rev.*, 2000, **100**, 143–167; (*b*) K. E. Riley, M. Pitonak, P. Jurecka and P. Hobza, *Chem. Rev.*, 2010, **110**, 5023–5063.
7 See for example: N. J. DeYonker, T. Grimes, S. Yockel, A. Dinescu, B. Mintz, T. R. Cundari and A. K. Wilson, *J. Chem. Phys.*, 2006, **125**, 104111-1–15 and references cited therein.
8 Y. Zhao and D. G. Truhlar, *Acc. Chem. Res.*, 2008, **41**, 157–167.
9 T. Schwabe and S. Grimme, *Acc. Chem. Res.*, 2008, **41**, 569–579.
10 (*a*) P. Y. Ren and J. W. Ponder, *J. Comput. Chem.*, 2002, **23**, 1497–1506; (*b*) J. W. Ponder, C. J. Wu, P. Y. Ren, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. A. DiStasio, M. Head-Gordon, G. N. I. Clark, M. E. Johnson and T. Head-Gordon, *J. Phys. Chem. B*, 2010, **114**, 2549–2564.

11 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
12 Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215–241.
13 K. S. Thanthiriwatte, E. G. Hehenstein, L. A. Brurns and C. D. Sherrill, *J. Chem. Theory Comput.*, 2011, **7**, 88–96.
14 S. Halbert, C. Clavaguéra and G. Bouchoux, *J. Comput. Chem.*, 2011, **32**, 1550–1560.
15 (*a*) J. W. Ochterski, G. A. Petersson and J. A. Montgomery, *J. Chem. Phys.*, 1996, **104**, 2598–2619; (*b*) J. A. Montgomery, M. J. Frish, J. W. Ochterski and G. A. Petersson, *J. Chem. Phys.*, 1999, **110**, 2822–2827; (*c*) J. A. Montgomery, M. J. Frish, J. W. Ochterski and G. A. Petersson, *J. Chem. Phys.*, 2000, **112**, 6532–6542.
16 (*a*) A. Baboul, L. A. Curtiss, P. C. Redfern and K. Raghavachari, *J. Chem. Phys.*, 1999, **110**, 7650–7657; (*b*) L. A. Curtiss and K. Raghavachari, *Theor. Chem. Acc.*, 2002, **108**, 61–70.
17 L. A. Curtiss, P. C. Redfern and K. Raghavachari, *J. Chem. Phys.*, 2007, **127**, 124105-1–8.
18 L. A. Curtiss, P. C. Redfern and K. Raghavachari, *J. Chem. Phys.*, 2007, **126**, 084108-1–12.
19 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09, Revision B.01*, Gaussian, Inc., Wallingford CT, 2010.
20 D. Jacquemin, C. Michaux, E. A. Perpète and G. Frison, *J. Phys. Chem. B*, 2011, **115**, 3604–3613.
21 J. W. Ponder, TINKER: Software Tools for Molecular Design, version 5.0, 2008, http://dasher.wustl.edu/tinker.
22 (*a*) J. Kaminsky and F. Jensen, *J. Chem. Theory Comput.*, 2007, **3**, 1774–1788; (*b*) D. Semrouni, G. Ohanessian and C. Clavaguéra, *Phys. Chem. Chem. Phys.*, 2010, **12**, 3450–3462.
23 G. Bouchoux, S. Huang and B. S. Inda, *Phys. Chem. Chem. Phys.*, 2011, **13**, 651–668.
24 S. Gronert and R. A. J. O'Hair, *J. Am. Chem. Soc.*, 1995, **117**, 2071–2081.
25 B. Lakard, *THEOCHEM*, 2004, **681**, 183–189.
26 M. Noguera, L. Rodriguez-Santiago, M. Sodupe and J. Bertran, *THEOCHEM*, 2001, **537**, 307–318.
27 (*a*) B. Lambie, R. Ramaekers and G. Maes, *J. Phys. Chem. A*, 2004, **108**, 10426–10433; (*b*) S. Blanco, M. E. Sanz, J. C. Lopez and J. L. Alonso, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **51**, 20183–20188.
28 R. Miao, C. Jin, G. Yang, J. Hong, C. Zhao and L. Zhu, *J. Phys. Chem. A*, 2005, **109**, 2340–2349.
29 M. Pecul, *Chem. Phys. Lett.*, 2006, **418**, 1–10.
30 T. C. Dinadayalane, G. N. Sastry and J. Leszczynski, *Int. J. Quantum Chem.*, 2006, **106**, 2920–2933.
31 C. M. Jones, M. Bernier, E. Carson, K. E. Colyer, R. Metz, E. Wischow, I. Webb, A. J. Andriole and J. C. Poutsma, *Int. J. Mass Spectrom.*, 2007, **267**, 54–62.
32 C. Bleiholder, S. Suhai and B. Paizs, *J. Am. Soc. Mass Spectrom.*, 2006, **17**, 1275–1281.
33 S. Gronert, D. C. Simpson and K. M. Conner, *J. Am. Soc. Mass Spectrom.*, 2009, **20**, 2116–2123.
34 L. F. Pacios, O. Galvez and P. C. Gomez, *J. Phys. Chem. A*, 2001, **105**, 5232–5241.
35 R. Wu and T. B. McMahon, *ChemPhysChem*, 2008, **9**, 2826–2835.
36 J. Oomens, J. D. Steill and B. Redlich, *J. Am. Chem. Soc.*, 2009, **131**, 4310–4319.
37 (*a*) M. Zhang and Z. Lin, *THEOCHEM*, 2006, **760**, 159–166; (*b*) V. Feyer, O. Plekan, R. Richter, M. Coreno, K. C. Prince and A. Carravetta, *J. Phys. Chem. A*, 2008, **112**, 7806–7815.
38 T. Szidarovsky, G. Czako and A. Csaszar, *Mol. Phys.*, 2009, **107**, 761–775.
39 X. Xu and Z. Lin, *THEOCHEM*, 2010, **962**, 23–32.
40 J. L. Alonso, C. Perez, M. E. Sanz, J. C. Lopez and S. Blanco, *Phys. Chem. Chem. Phys.*, 2009, **11**, 617–627.
41 A. Fernandez-Ramos, E. Cabaleiro-Lago, J. M. Hermida-Ramon, E. Martinez-Nunez and A. Pena-Gallego, *THEOCHEM*, 2000, **498**, 191–200.
42 (*a*) J. C. Dobrowolski, J. E. Rode and J. Sadlej, *THEOCHEM*, 2007, **810**, 129–134; (*b*) S. M. Bachrach, T. T. Nguyen and D. W. Demoin, *J. Phys. Chem. A*, 2009, **113**, 6172–6181.
43 Z. Tian, A. Pawlow, J. C. Poutsma and S. R. Kass, *J. Am. Chem. Soc.*, 2007, **129**, 5403–5407.
44 M. E. Sanz, S. Blanco, J. C. Lopez and J. L. Alonso, *Angew. Chem., Int. Ed.*, 2008, **47**, 6216–6220.
45 J. J. Wilke, M. C. Lind, H. F. Schaefer III, A. G. Csaszar and W. D. Allen, *J. Chem. Theory Comput.*, 2009, **5**, 1511–1523.
46 H. K. Woo, K. C. Lau, X. B. Wang and L. S. Wang, *J. Phys. Chem. A*, 2006, **110**, 12603–12606.
47 R. Linder, K. Seefeld, A. Vavra and K. Leinermanns, *Chem. Phys. Lett.*, 2008, **453**, 1–6.
48 O. Plekan, V. Feyer, R. Richter, M. Coreno, M. de Simone, K. C. Prince and V. Carravetta, *J. Phys. Chem. A*, 2007, **111**, 10998–11005.
49 H. Lioe, R. A. J. O'Hair, S. Gronert, A. Austin and G. E. Reid, *Int. J. Mass Spectrom.*, 2007, **267**, 220–232.
50 S. Desaphy, C. Malosse and G. Bouchoux, *J. Mass Spectrom.*, 2008, **43**, 116–125.
51 P. B. Armentrout, A. Gabriel and R. M. Moision, *Int. J. Mass Spectrom.*, 2009, **283**, 56–68.
52 (*a*) M. J. Locke, R. L. Hunter and R. T. McIver Jr., *J. Am. Chem. Soc.*, 1979, **101**, 272–273; (*b*) M. J. Locke and R. T. McIver Jr., *J. Am. Chem. Soc.*, 1983, **105**, 4226–4232; (*c*) M. J. Locke, *Ph.D. Thesis*, University of California, Irvine, 1981.
53 E. P. Hunter and S. G. Lias, *J. Phys. Chem. Ref. Data*, 1998, **27**, 413–656.
54 G. Bouchoux and J. Y. Salpin, *Eur. J. Mass Spectrom.*, 2003, **9**, 391–402.
55 R. A. J. O'Hair, J. H. Bowie and S. Gronert, *Int. J. Mass Spectrom. Ion Processes*, 1992, **117**, 23–36.
56 J. E. Bartmess, *NIST Chemistry WebBook*, in *NIST Standard Reference Database Number 69*, ed. W. G. Mallard and P. J. Linstrom, National Institute of Standard and Technology, Gaithersburg, MD, 2008.
57 X. Li and A. G. Harrison, *Org. Mass Spectrom.*, 1993, **28**, 366–371.
58 G. Bojesen and T. Breindahl, *J. Chem. Soc., Perkin Trans. 2*, 1994, 1029–1037.
59 C. Afonso, F. Modeste, P. Breton, F. Fournier and J. C. Tabet, *Eur. J. Mass Spectrom.*, 2000, **6**, 443–449.
60 (*a*) A. G. Harrison, *Mass Spectrom. Rev.*, 1997, **16**, 201–217; (*b*) M. K. Green and C. B. Lebrilla, *Mass Spectrom. Rev.*, 1997, **16**, 53–71.
61 R. J. Nieckarz, C. G. Atkins and T. B. McMahon, *ChemPhysChem*, 2008, **9**, 2816–2825.

